MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS 1963-A

AFGL-TR-84-0084

AD-A142 159

ANALYSIS LINKING THE TENSOR STRUCTURE
TO THE LEAST-SQUARES METHOD

Georges Blaha

Nova University Oceanographic Center
8000 North Ocean Drive
Dania, Florida 33004

Scientific Report No. 3

January 1984

DTIC
SELECTED
JUN 1 9 1984
A

84 06 18 067
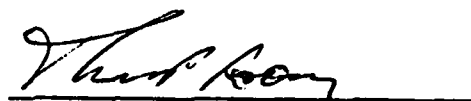
CONTRACTOR REPORTS

This report has been reviewed by the ESD Public Affairs Office (PA) and is releasable to the National Technical Information Service (NTIS).

This technical report has been reviewed and is approved for publication.

GEORGE HADGIGEORGE
Contract Manager

THOMAS P. ROONEY
Chief, Geodesy & Gravity Branch

FOR THE COMMANDER

DONALD H. ECKHARDT
Director
Earth Sciences Division

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFGL-TR-34-0084	2. GOVT ACCESSION NO. AD A142159	3. RECIPIENT'S CATALOG NUMBER
4. TITLE *(and Subtitle)* ANALYSIS LINKING THE TENSOR STRUCTURE TO THE LEAST-SQUARES METHOD		5. TYPE OF REPORT & PERIOD COVERED Scientific Report No. 3
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Georges Blaha		8. CONTRACT OR GRANT NUMBER(s) F19628-82-K-0007
9. PERFORMING ORGANIZATION NAME AND ADDRESS Nova University Oceanographic Center 8000 North Ocean Drive Dania, Florida 33004		10. PROGRAM ELEMENT. PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2309G1BB
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Geophysics Laboratory Hanscom AFB, Massachusetts 01731 Contract Monitor: George Hadgigeorge/LWG		12. REPORT DATE January 1984
		13. NUMBER OF PAGES 105
14. MONITORING AGENCY NAME & ADDRESS(*if different from Controlling Office*)		15. SECURITY CLASS. *(of this report)* Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

Least-squares adjustment	Tensor analysis	Contravariant
Parametric method	Metric tensor	Intrinsic
Condition method	Associated metric tensor	Extrinsic
Variance-covariances	Covariant	Contraction

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

One of the main purposes of the geometrical approach to the least-squares adjustment as presented herein is to describe the adjustment theory in a simple and plausible way and, at the same time, to establish a niche for such an approach in a field that has been explored decades ago and then again in recent years. This development is based on differential geometry with tensor structure and notations. In expressing the desired tensor relations, it relies heavily on orthonormal space and surface vectors and on the extrinsic properties of —

DD $_{1\ JAN\ 73}^{FORM}$ 1473 EDITION OF 1 NOV 65 IS OBSOLETE

i

Unclassified

surfaces linking the two kinds of vectors. In order to relate geometry to adjustments, the geometrical concepts are extended to an n-dimensional space and u-dimensional or r-dimensional surfaces, where n is the number of observations, u is the number of parameters in the parametric (or observation equation) method and r is the number of condition equations in the condition method, with n=u+r. Other methods are not treated here.

The adjustment of the above two methods is associated with one and the same geometrical situation, in which a given observational vector is decomposed into a model vector lying in the (u-dimensional) model surface and a second vector orthogonal to this surface. The weight matrices are made to correspond to metric tensors or, in an equivalent approach, to associated metric tensors. The adjustment formulas for either method follow from a direct transcription of the corresponding geometrical relationships. Whether expressing the geometrical or least-squares relationships, the diagrammatic representation is adopted. Only one diagram is needed for the description of both adjustment methods. Connection is made to Hilbert spaces by demonstrating that the tensor approach to the least-squares adjustment is a classical case of the Hilbert-space approach.

Accession For

NTIS GRA&I
DTIC TAB
Unannounced
Justification

By
Distribution /

Availability

Dist

A1

TABLE OF CONTENTS

LIST OF FIGURES

1. INTRODUCTION AND HISTORICAL NOTES

The goal of the present report is the development of the least-squares theory and results in a purely geometrical manner. Because of possible linearization of the adjustment models involved, the term "geometry" can be understood as "differential geometry" in this context. The tool adopted in perceiving the adjustment procedure through the eyes of geometry is tensor analysis, although this preference may have been influenced by the author's appreciation of the tensor approach as introduced to theoretical geodesy by Antonio Marussi, [1], and elaborated on by the late Martin Hotine, [2].

Little would be accomplished if the known adjustment formulas should only be translated into tensor notations or if they should be rederived following an old route with new notations. Instead, new geometrical concepts will be introduced into the derivations in order to elucidate the least-squares (L.S.) process and make its understanding compact. The derivation of all the adjustment formulas needed in the parametric method (also called the observation equation method) and in the condition method will be accomplished simply by identifying certain tensors in a purely geometrical situation with the vectors and matrices given in one or the other adjustment setup.

In the closing paragraphs the geometrical approach will also serve to illustrate the Hilbert-space approach to the L.S. adjustment. For example, the contravariant and covariant tensor "spaces" can be identified with dual Hilbert spaces, the scalar product of two vectors in tensor notations can be identified with the same product in the Hilbert space, etc., so that the

tensor formulation becomes in fact a classical Hilbert-space formulation. This idea was inspired by Petr Vaníček's paper, "Diagrammatic Approach to Least Squares", prepared for a public seminar at the University of Stuttgart, August 19, 1982.

In order to acquaint himself with the basic premise of the present geometrical approach, the reader can imagine that a certain "observational vector" dx having contravariant components dx^r is decomposed into two vectors having contravariant components dx'^r and dx''^r, the first lying in a "model surface" and the second completing the system of equations $dx^r = dx'^r + dx''^r$ (in [2] this is called "vector equation"). The lengths of these three vectors, denoted respectively as ds, ds' and ds", are tensor invariants, i.e., quantities independent of the man-made coordinate system. If the geometrical quantity ds''^2 is identified with the L.S. quadratic form (often denoted $V^T PV$) required to fulfill the familiar minimum condition, it follows in the geometrical context that dx" must be orthogonal to the model surface regardless of the mathematical formulation of the surface, the nature of the space and surface coordinates, etc.

Upon extending this concept to an n-dimensional space and a u-dimensional (hyper-) surface, dx" is seen to lie in the r-dimensional subspace orthogonal to the model surface at a given point, where $r = n - u$. This subspace will be called "second surface" for convenience. Clearly, in a three-dimensional space such a "surface" would reduce to a straight line perpendicular to the model surface; if the model "surface" were a straight line, the second surface would be a plane perpendicular to it, upon considering a small neighborhood of the point in question as implied in this study.

-2-

Admittedly, the tensor mechanism is not the only tool leading to the L.S. estimates via geometry. For example, the generalized Hilbert-space mechanism is no less "geometrical". The avenue one chooses to pursue depends to a great extent on personal preference and background. In the generalized Hilbert-space approach as used by the "Fredericton school" (to be mentioned later in this chapter), a coordinate system is implied through the Riemannian metric but it is not used explicitly. Indeed, an explicit coordinate system is not necessary in representing the geometrical relations which are invariant under coordinate transformations.

However, the coordinate systems as used in the present geometrical approach introduce neither limitations nor "excess baggage". The first assertion follows from the fact that no particular coordinate system is required by the derivations, only the existence of space and surface coordinate systems is assumed. And the second assertion becomes apparent upon the realization that the "coordinates" are linked directly to the desired L.S. quantities. For example, the space coordinates x^r are used in representing the observations, the surface coordinates u^α implying the model surface are used in representing the adjusted parameters, etc. In other approaches such quantities have to be represented as well, only the means of their identification vary.

The technique which is the cornerstone of this presentation consists in the "vectorization" of tensors. It is well known that a space tensor, such as the second-order metric tensor, can be expressed by means of vectors (first-order tensors) belonging to the same point in space. However, a set of partial derivatives $\partial x^r / \partial u^\alpha$ relating the space coordinates x^r and the

surface coordinates u^α at a given point on the surface is also of
fundamental importance in tensor analysis. According to §3 and §12 in
Chapter 6 of [2], it is a tensor. In particular, it transforms like the
contravariant space components and the covariant surface components of a
vector lying in the surface. The latter property can be illustrated by
considering each space coordinate as a scalar defined over the surface,
in which case the corresponding quantities in the set $\partial x^r/\partial u^\alpha$ form the
surface gradient of a scalar.

The tensor character of $\partial x^r/\partial u^\alpha$ implies that it can be subject to
covariant differentiation. Indeed, the surface covariant derivative of
this tensor is particularly useful in differential geometry. It is explained
in §19 on page 34 of [2] and used, for example, on the subsequent pages of
[2] containing the development of the standard results, such as the Gauss
equations of the surface, the Weingarten equations, the Mainardi-Codazzi
equations, or the Gaussian curvature of the surface. However, the tensor
character and, especially, the covariant derivatives of $\partial x^r/\partial u^\alpha$ are not
exploited in this presentation. The point which should be emphasized is
that the quantities such as $\partial x^r/\partial u^\alpha$ can be vectorized.

It will be of great importance that $\partial x^r/\partial u^\alpha$ can be expressed by means
of orthonormal vectors lying in the surface and viewed in their space as
well as surface context (i.e., considered in both space and surface
coordinates). In order to find such a vectorized expression it is neces-
sary to make use of the extrinsic properties of the surface which relate
it to the surrounding space, in addition to the intrinsic properties
concerned with this surface as a space of certain dimensions in its own

-4-

right. Accordingly, several properties of u- and r-dimensional surfaces embedded in an n-dimensional space will be developed, representing a generalization and an extension of the properties encountered in [2], the first part of Chapter 6, where $u = 2$ and $n = 3$.

The fundamental role in the geometrical development via the vectorization technique is played by a set of orthonormal vectors divided into those lying in the model surface and the remainder of the set (the latter are thus lying in the second surface according to the earlier definition). These vectors can serve to produce most of the L.S. results and relatior ships in a simple and straightforward manner. In particular, adjustmen formulas can be arrived at without algebraic operations such as the di ferentiation of the L.S. quadratic form with respect to the parameters ι. the introduction of the Lagrange multipliers, etc.

It should be pointed out that this geometrical approach is concerned with the derivation of the L.S. estimates as well as of the corresponding variance-covariances and weights, but does not attempt to draw any probabilistic implications. Accordingly, no assumptions (such as normality) are needed with regard to the error probability distribution in the observations. The initial geometrical setup is thus consistent with the "geometrical interpretation" of Section 1.3 in [13] which will be briefly discussed later in conjunction with Fig. 1.

The property allowing a very natural derivation of the L.S. results via geometry -- not merely their geometrical interpretation -- is the complete correspondence between the variance-covariance propagation on one

hand and the "propagation" of the associated (contravariant) metric tensors on the other hand. In the same context, the "propagation" of the (covariant) metric tensors describes what could be termed the weight propagation. This correspondence will eventually lead to new expressions and relationships, such as the formulas giving the (singular) weight matrices for the adjusted observations and for the residuals.

The correspondence between the weight matrix of the observations and the space metric tensor is the key element allowing the earlier mentioned representation of the L.S. quadratic form by the geometric quantity $ds^{"2}$. The adjustment problem can thereby be transformed into the geometrical problem which will be formulated as the minimization of the length of the residual vector in the space metricized by the weight matrix of the observations. The geometric derivation of the L.S. results will then follow upon representing the remaining quantities of the L.S. setup by the appropriate geometric quantities. It will also become apparent that the same L.S. results can be obtained with the covariant and contravariant indices interchanged.

In the historical context, this study can be placed somewhere between the "Fredericton school" represented by the works of Vaníček ([3] and the earlier mentioned Seminar), Mohammad-Karim ([4]) and others, and the "Delft school" represented, over a period spanning more than two decades, by Tienstra ([5], [6], [7]), Baarda ([8], [9], [10]), Kooimans ([11]) and a number of others (for example, the editors of [7]). It would probably be more appropriate to place these two schools and the present study at the

vertices of a triangle, due to significant differences in mathematical tools, approach and scope. The Delft school is noted for its algebraic approach using the principles and notations of tensor analysis, as well as for the breadth and scope of its extensive publications. The Fredericton school is noted for its diagrammatic approach to the representation of various L.S. methods and, as in the Seminar, for its Hilbert-space approach to the L.S. adjustment. Finally, this report emphasizes the geometrical approach with tensor structure to the L.S. adjustment theory. Although its scope is limited to the parametric and condition methods, the current approach can be expanded in the future to include other L.S. methods and topics.

The trends followed by the above two schools are now briefly summarized. With regard to the publications of the Fredericton school, the subject of [3] is limited to the parametric method which is also the most extensively treated topic in the Seminar. However, the latter presents two additional adjustment models as well (the condition method and the general method involving several or all residuals in each equation). But the most prominent feature that distinguishes the Seminar from [3] is the emphasis on Hilbert-space interpretation. In particular, whereas [3] associates certain tensor quantities with the corresponding vectors and matrices of the L.S. setup and proceeds to a commutative diagram, the Seminar also relates the tensor operations to those in the Hilbert space. The reference [4] deserves mentioning for its treatment of several adjustment models including collocation. It builds on the ideas put forth in [3] of which it could be considered a natural extension. With regard to the parametric method, only a

few unessential features, such as an iterative scheme, distinguish it from
[3]. In both [3] and [4] the weight matrices (P_ℓ for the observations and
P_x for the parameters) are represented by the associated metric tensors,
see e.g. the statement on page 222 of [3] that "the metric tensors P_ℓ and
P_x are considered contravariant."

In an earlier version of the present study the weight matrices, called
here P-matrices, were identified exclusively with the metric tensors (i.e.,
treated as covariant), whereas the variance-covariance matrices, called
here C-matrices, were identified exclusively with the associated metric
tensors (i.e., treated as contravariant). This was due especially to a
simple and convenient geometrical interpretation of the design matrix in
the parametric method. Since the conception of the idea linking certain
tensor notations and operations to the L.S. adjustment and mathematical
statistics, the Delft school has adhered strictly to this kind of identifi-
cation (with the exception of a scale factor) as is evidenced by [5] through
[11], hence the term "traditional identification". Although in [3]
Vaníček broke with this tradition by reversing the covariant and contra-
variant indices in what will be called "new identification", in the Seminar
he reverted back to it. It is apparent that by trying to proceed in both
ways, Vaníček anticipated a perfect duality between the covariant and contra-
variant "spaces". Such a duality points to a close parallelism between the
tensor approach and the Hilbert-space approach to the adjustment theory
alluded to at the beginning of this chapter.

Indeed, when choosing the traditional identification of the given
P-matrix in the Seminar, Vaníček stated that whether P is "considered to

-8-

correspond to covariant or contravariant metric tensor is a more or less arbitrary decision". The present study fully supports this assertion in an evaluation which is essentially a by-product of the vectorizations mentioned earlier. In such a process all the pertinent relationships are expressed as tensor equations since difficulties would arise if any dependence on coordinate systems were introduced.

As will become clear, passing from the traditional to the new identification entails either changing the rules for matrix multiplications or transposing the matrices as they are substituted for the corresponding second-order tensors. The latter is preferred by far. This interchangeability in identification will be demonstrated not only directly in tensor form (the "weight propagation law" will be again seen to follow exactly the same rules as the known variance-covariance propagation law), but also by carrying out the L.S. adjustment in terms of the parametric and condition methods with both kinds of identification. All the results between the two versions will be shown to be identical and the corresponding commutative diagrams will be seen to be mirror images of each other.

In the traditional identification the design matrix (A) of the parametric method is associated with the partial derivatives $\partial x^r/\partial u^\alpha$ as introduced earlier, with x^r representing the space coordinates and u^α representing the model-surface coordinates. This association was originally inspired by [3]. The geometrical pursuit of this concept leads to expressing the same "model vector" both in space coordinates, in which case its contravariant components are written as dx'^r and correspond essentially to the adjusted observations, and in model-surface coordinates, in which case its

contravariant components are written as du^α and correspond essentially to the adjusted parameters. The appropriate second-order metric tensors (g'_{rs} and $a_{\alpha\beta}$ for the P-matrices, g'^{rs} and $a^{\alpha\beta}$ for the C-matrices) then follow from the geometrical setup. Most of the geometrical insight offered thus far has been linked to the traditional identification, in which the interpretation of the matrix A is particularly simple and natural. In the new identification the covariant and contravariant indices would be interchanged with regard to the P- and C-matrices as well as with regard to various adjustment vectors; furthermore, the interpretation of the matrix A would undergo a basic change.

Reasoning similar to the above, but with regard to dx" and the second surface, leads to the (minus) residuals and the "adjusted parameters" for the condition method as well as to the corresponding P- and C-matrices. It should be noted that these "parameters" arise from purely geometrical considerations and are linked to the other geometrical quantities in relationships reminiscent of the parametric method; the different kinds of final formulas stem merely from the different a priori design matrices. It should also be noted that such "parameters" are not the Lagrange multipliers known from the standard treatment of the condition method, although the two are closely related and, in tensor notations, differ only in sign and by having the contravariant and covariant indices interchanged. Clearly, the Lagrange multipliers result from algebraic operations and are not needed in the present approach. Nor do they lend themselves readily to a geometrical interpretation.

The commutative diagrams, although inspired by [3], are used here on a

broader basis than has been the case with the Fredericton school. In
particular, not only the final product (the adjustment scheme in terms of
vectors and matrices) is expressed in this way, but also the underlying
geometrical representation (the relationships in terms of vectors and
second-order tensors). In another extension to the Fredericton approach,
the special attention paid to the model vector dx' lying in the model sur-
face and to the vector dx'' lying in the second surface leads to a new out-
come giving the corresponding singular P- and C-matrices as distinguished
from the nonsingular P- and C-matrices associated with the observational
vector dx. In the standard adjustment theory these two new C-matrices are
known as the C-matrices for the adjusted observations and the residuals,
respectively, but the corresponding P-matrices are not considered. Still
another addition to the adjustment theory of the Fredericton school is the
geometrical derivation of the C-matrices for linear functions of adjusted
parameters or adjusted observations.

A brief review of highlights in the Delft approach is in order,
especially for comparison purposes between the references [5] - [11] and
the present study. The difference in scope has been already mentioned, as
well as the adherence of the Delft school only to the traditional identifica-
tion. It has also been indicated that the Delft approach is basically
algebraic, an extensive use of tensor notations notwithstanding. This can
be noticed already in the differentiation of the L.S. quadratic form with
respect to the parameters in the parametric method or in the introduction
of the Lagrange multipliers in the condition method. Although [5] - [11]
contain certain geometric illustrations, one of the main differences between

the Delft approach and the geometrical approach pursued herein is that the former does not exploit such concepts as the extrinsic properties of surfaces. A corroboration of this statement will be made in conjunction with [11].

Reference [5]. This reference lays out the foundation of the Delft approach, encompassing the traditional identification, the tensor notations, etc. The derivations on the beginning pages show how the contravariant components of vectors and the associated metric tensors are related in linear transformations. The adjustment formulas are derived with uncorrelated observations and then with correlated observations, where the correlations are introduced by a linear transformation of the originally uncorrelated variates. The condition method is treated first, followed by the parametric method. The use of the Lagrange multipliers and the differentiation of the L.S. quadratic form with respect to the parameters form an essential part of [5]. However, this has been already indicated in the preceding paragraph and need not be repeated in the balance of the Delft approach summary. The latter part of [5] treats the parametric method with constraints and the general method based on the formulas already derived; however, these topics go beyond the scope of the present study.

Reference [6]. Here the emphasis is put on the statistical concepts such as theory of observations, distributions (normal, multivariate-normal), correlation, residuals, etc. The L.S. properties are studied against the background of the normal distribution of observations. A geometric interpretation given in conjunction with the condition method views the observations as the n coordinates of the "observed point". It should be pointed

-12-

out that a similar philosophy is used, and elaborated upon, also in the present report through the introduction of the observational vector dx with n components. The solution for the condition method in [6] is presented in a standard form as well as in a form characterized by a unit weight matrix as obtained from the original matrix through a coordinate transformation.

One may notice that the use of concepts based on true errors, true values, etc., is discouraged in [6]. Instead, the case is made in favor of the physical reality of the observations. The present study could easily avoid any mention of true values, etc., especially since it is not statistically oriented. However, in a few isolated instances the term "true" is used for illustration. Such a designation should be understood in a loose sense, mainly as serving to describe certain relationships in a given adjustment model. Although such "true" quantities will never be known, their estimates -- which represent the outcome and the real purpose of the adjustment -- have to satisfy the same relationships. One can thus regard the mathematical formulation of an adjustment model in terms of "true" values merely as a matter of convenience.

Reference [7]. In this reference the Delft approach is given a detailed mathematical foundation. The first three chapters expand parts of the theory outlined in [6], such as frequency distribution, normality characteristics of observations, or multivariate-normal distribution free of correlation. In Chapter 4 the correlations are introduced by linear transformations of uncorrelated variates, similar to [5]. Chapter 5 proceeds to an adjustment

-13-

of the condition method under the normality assumption, again in line with [5]. In Chapters 6 and 7 computational algorithms are developed consistent with the previous derivations, especially the algorithms for the condition and parametric methods with uncorrelated and correlated observations. In addition, the parametric method with constraints, the general method, and the general method with constraints are treated; these topics are beyond the scope of the present report, as are the topics developed in Chapters 8 and 9 dealing, for example, with the analysis of residuals.

Reference [8]. The initial steps in this reference are based on the previous concepts established especially in [7]. The mathematical symbolism is characterized by mixed tensor and matrix notations. Although the normal distribution of observations is accepted for simplicity, other distributions are not excluded. The previous work on adjustment in steps as well as on statistical treatment of adjustment results and on other topics beyond the scope of the present study is expanded, including testing of hypotheses, non-linearity effects, etc. In spite of the differences between the Delft approach and the geometrical approach, the formulas that can be compared show a complete agreement. For example, Mr. P. Teunissen of the Technical University of Delft showed, after some unessential notational changes, a number of equivalences existing between this study and [8].

Reference [9]. In many respects this reference is an expansion and continuation of [8]. Accordingly, the previous notations and conventions are preserved. It is concerned with testing of hypotheses (null hypothesis, one or more alternative hypotheses), analysis of variance, confidence regions, reliability of geodetic networks, etc. Such statistical treatment

-14-

of the adjustment outcome is not a part of the present report.

Reference [10]. This reference, available only in Dutch, appears to
have been conceived as a textbook (in two volumes) addressing itself to
readers sufficiently familiar with statistical concepts in the adjustment
theory as well as with tensor notations and operations. It is based on
some of the material contained in [5] - [9] and proceeds to a substantial
expansion of this approach. For example, several initial formulas as well
as the notations correspond to Chapter 2 of [9]; the presentation of the
solution for the parametric and condition methods in Chapter 7 resembles
the first part of Chapter 2 in [8]. The theory for these two methods is
recapitulated and broadened, including linear transformations, rank of
matrices, etc. The expanded and newly treated concepts (always in tensor
notations) going beyond the scope of the present study include ortho-
gonalization, application of the Choleski algorithm, adjustment in phases,
treatment of the parametric method with constraints as well as the general
method with and without constraints, treatment of non-linear models, etc.

Reference [11]. This reference, although dating back to 1958, is
left for the end of the list in order to illustrate, in concrete terms,
the most noticeable difference between the Delft approach and the geometrical
approach introduced in this report. The scope of [11] is limited to re-
capitulating some of the principal concepts in [5] - [7], as well as to
developing the condition method along those lines including a certain amount
of testing. The basic formulas for the parametric method are obtained upon
adopting a different mathematical form for the condition method. The point

-15-

of interest lies here in the transformation formula for the associated
metric tensor. With slight changes in notations, this reference utilizes

$$\bar{g}^{ab} = (\partial \bar{x}^a/\partial x^i)(\partial \bar{x}^b/\partial x^j)g^{ij} \, ,$$

which is also one of the basic formulas in the present study, expressing
the associated metric tensor in $\bar{x}^r$ coordinates based on the corresponding
tensor in x^r coordinates; all the index letters pertain to an n-dimensional
space (i.e., they run from 1 to n).

In the adjustment context the above formula expresses the law of
propagation of variance-covariances. The present approach attributes to
this formula the same meaning as [11] in the category of the coordinate
transformations (from x^r to $\bar{x}^r$ coordinates) in one and the same n-dimensional
space. If more than one space is considered, the first (lower-dimensional)
being immersed in another (higher-dimensional), the above law can be extended.
Indeed, this avenue is followed herein upon identifying the former space
with a surface embedded in the latter. The formulas for the desired results
including variance-covariances and weights can then be obtained by utilizing
the extrinsic properties of surfaces and surface vectors instead of utilizing
the algebraic means as in the Delft approach.

The orientation of the present study in the historical context having
been established, it is useful to state clearly its limitations and summarize
its special features. With regard to the limitations, only the parametric
and condition methods are presented, with no consideration given to the
general method (with or without constraints) or collocation; the design

-16-

matrix of the parametric method is considered to have the full column rank

so that the rank deficient systems usually associated with the adjustment

of free networks are not treated either. Another limitation concerns the

statistical concepts which are represented here only by the formulas giving

the variance-covariances and the corresponding weights; no consideration

is given to probability distributions of random variables, confidence

regions, testing of hypotheses, etc. Finally, no attempt is made to treat

nonlinear problems (here the power of tensor calculus could be particularly

useful). These three kinds of limitations indicate the topics which could

be addressed in the future by other studies, reports or papers based on the

geometrical approach. For example, properties of the point functions g_{rs},

etc., could be exploited in a nonlinear (and non-linearized) L.S. problem

wherein the point to which these functions belong could be subject to dis-

placements.

The highlights of this study can be summarized as follows:

a) Differential geometry based on the tensor structure and notations

is used throughout; advantage is taken of the simplicity in expressing the

pertinent tensor quantities by means of orthonormal space and/or surface

vectors.

b) The geometrical concept is generalized to an n-dimensional space

and u-dimensional or r-dimensional surfaces; n is identified with the number

of observations, u with the number of parameters in the parametric method

and r with the number of condition equations in the condition method, where

$n = u + r$.

c) The possibility of identifying the associated metric tensors (contravariant) with C-matrices and the metric tensors (covariant) with P-matrices is established, as is the possibility of the reverse identification; the first identification is called traditional and the second is called new.

d) Adjustment in terms of the parametric and condition methods is associated with one and the same geometrical situation, in which a given observational vector is decomposed into a model vector lying in the (u-dimensional) model surface and a second vector orthogonal to this surface; the adjustment formulas for either method follow from a direct transcription of the corresponding geometrical relationships.

e) In expressing the geometrical relationships as well as their L.S. counterparts, here the parametric and condition methods, the diagrammatic representation is adopted; only one diagram is needed for the description of both methods.

f) The parametric and condition methods are expressed in either type of identification (traditional and new); the corresponding formulas are identical and the corresponding diagrams are mirror images of each other, implying duality between the two identifications similar to the duality properties in Hilbert spaces.

g) Connection to Hilbert spaces is established by demonstrating that the tensor approach to the L.S. adjustment is a classical case of a Hilbert-space approach to the same problem; this could help to elucidate the Hilbert-space theory for those who are familiar with a few basic principles of tensor analysis.

-18-

One of the main purposes of the geometrical development presented herein is to describe the adjustment theory in a simple and plausible way. The tools for carrying out this task have been reduced to a minimum, consisting of tensor relations contained in the beginning sections. Yet they are sufficient to make the derivations completely general. However, the tensor approach alone cannot guarantee simplicity. For example, failing to proceed via vectorization could complicate and significantly lengthen the development. The present report is intended to be entirely self-contained. The references are not needed for the derivations, their main role is to establish a niche for the geometrical approach to the L.S. adjustment in a field that has been explored decades ago and then again in recent years.

2. MATHEMATICAL BACKGROUND

This chapter is divided into two sections, the first describing the space relations and the second introducing surfaces and surface vectors used for the derivation of the L.S. formulas. All the relations are considered at a given point P. (not to be confused with P-matrices). The n-dimensional space is spanned by n orthonormal vectors which can be considered its base vectors. The first u vectors are denoted by the Roman letters ℓ, j, etc., while the last $r = n - u$ vectors are denoted by the Greek letters ν, etc. In either group dots will take place of the vectors not written explicitly. The first u vectors are chosen to lie in the u-dimensional model surface. Therefore, they can also be expressed in surface coordinates and considered as orthonormal base vectors in such a u-dimensional subspace. The space components of tensors (including vectors) are designated by the lower-case Roman letters as indices. The surface components are designated by the Greek letters as indices. Eventually, the r-dimensional subspace orthogonal to the model surface at P will be identified with a second surface and the upper-case Roman letters as indices will be resorted to.

Most notations will be used consistent with [2]. Accordingly, the space metric tensor is represented by g_{rs} (this notation also indicates its r,s components) and the metric tensor for the model surface by $a_{\alpha\beta}$; the metric tensor for the second surface will be written in the form a_{LM}. The associated metric tensors have both indices raised. The basic notions of tensor algebra need not be repeated since they can be found in standard

textbooks, e.g., in the beginning chapters of [2]. Thus, the summation convention for repeated (dummy) indices applies, a differential vector divided by its length is a unit vector which may be identified by its contravariant or covariant components and may be expressed in space or surface coordinates (for example, $\ell^r = dx'^r/ds'$ are the contravariant space components of a unit vector ℓ lying in the model surface), all the metric tensors are defined as symmetrical, etc. In addition to second-order covariant tensors, such as g_{rs}, and contravariant tensors, such as g^{rs}, also second-order mixed tensors will be used, such as δ^r_s (it equals 0 for $r \neq s$ and 1 for $r = s$); the latter is known as the Kronecker delta and has similar properties in any space or surface coordinates.

Although the concepts exposed in parts of the first six chapters of [2] provide valuable background material, several formulas developed in this study do not have an equivalent in [2], and several others are derived in a different and independent manner. The two most noticeable differences that distinguish the derivations herein from those in [2] are the higher-dimensional approach to space and surface relationships (in [2], n=3 and u=2), and the choice to carry out the derivations in general coordinates rather than to base some of them on the results valid in Cartesian coordinates (this choice is just as easy, and it removes any possible doubts one may feel when relying on Cartesian coordinates with indistinguishable contravariant and covariant components, even though the relationships expressed in tensor equations are valid in any coordinates and make the use of the Cartesian system permissible).

A minor remark may be added regarding the terminology. In [2], the term "vector" is utilized to identify a physical entity (such as ℓ), as well as its contravariant components (ℓ^r) or covariant components (ℓ_s), see e.g. the statement on page 7 therein, "three mutually perpendicular unit vectors λ_r, μ_r, ν_r, ... ". In the present report a distinction is made with regard to the above three characteristics.

2.1 Basic Relations in Space

The differential coordinate changes in the x^r system based on such changes in the $\overline{x}^r$ system and vice versa constitute a simple starting point in the derivations:

$$dx^r = (\partial x^r/\partial \overline{x}^s)d\overline{x}^s \ , \qquad d\overline{x}^i = (\partial \overline{x}^i/\partial x^t)dx^t \ . \tag{1a,b}$$

Upon dividing both sides of (1a,b) by ds, the length of the vector dx, these formulas are seen to hold true with regard to unit vectors; and upon multiplying them further by a scalar they are seen to hold true with regard to arbitrary vectors as well. The formulas (1a,b) describe the transformation of the contravariant components of vectors in a space of arbitrary dimensions, such as n.

The square of ds for any vector dx is defined as

$$ds^2 = g_{rs}dx^r dx^s = \overline{g}_{ij}d\overline{x}^i d\overline{x}^j \ , \tag{2}$$

where the second equality stems from ds being an invariant. Upon replacing $d\overline{x}^i$, $d\overline{x}^j$ by the appropriate forms from (1b), we obtain

$$[g_{rs} - \overline{g}_{ij}(\partial \overline{x}^i/\partial x^r)(\partial \overline{x}^j/\partial x^s)]dx^r dx^s = 0 \ .$$

Since the metric tensor is symmetric, the expression inside the brackets can have up to n(n+1)/2 distinct components in an n-dimensional space. The product $dx^r dx^s$ represents n(n+1)/2 independent coefficients (dx in either factor is the same arbitrary vector) connecting in general the same number of distinct components inside the brackets as has just been shown. In

following the reasoning on page 11 of [2] for a general case (not only n=3), one concludes that the expression inside the brackets must be zero, and thus

$$g_{rs} = (\partial \bar{x}^i / \partial x^r)(\partial \bar{x}^j / \partial x^s) \bar{g}_{ij} \; . \tag{3}$$

Instead of expressing ds^2, (2) could have been used in conjunction with two independent vectors dx, dy to produce a quantity called the scalar product, also an invariant under coordinate transformation. Since the product $dx^r dy^s$ would now represent n^2 independent coefficients rather than $n(n+1)/2$, equation (3) would be arrived at even if g_{rs} were replaced by a general non-symmetric tensor h_{rs}. One can thus conclude that the condition of invariance gives rise to the transformation formula such as (3) for any second-order covariant tensor.

Next we define covariant components of an arbitrary vector according to

$$dx_s = g_{sr} dx^r \; , \tag{4}$$

valid in any coordinates, i.e., also with all three tensor quantities overbarred. According to (2) one can thus write

$$ds^2 = dx_s dx^s = d\bar{x}_s d\bar{x}^s \; . \tag{5}$$

Upon utilizing (1a) and (3) in (4) and taking advantage of the Kronecker delta in

$$(\partial \bar{x}^i / \partial x^r)(\partial x^r / \partial \bar{x}^t) = (\partial \bar{x}^i / \partial \bar{x}^t) = \bar{\delta}^i_t = \delta^i_t \; ,$$

after some rearranging of indices we obtain

-24-

$$dx_r = (\partial\bar{x}^s/\partial x^r)d\bar{x}_s \ , \qquad d\bar{x}_i = (\partial x^t/\partial\bar{x}^i)dx_t \ , \tag{6a,b}$$

where (6b) follows from (6a) upon interchanging the overbars and re-arranging the indices. The formulas (6a,b) describe the transformation of the covariant components of vectors. The comparison of (3) and (6a) confirms that the covariant tensor components transform like the covariant vector components for each index. As a manner of interest we observe that if (6a) is contracted with the contravariant components A^r of an arbitrary vector A (for the right-hand side these components are written in a form similar to the right-hand side of 1a), the resulting product is confirmed to be a tensor invariant. In [2], page 7, this invariance condition was used to derive the covariant transformation (6a,b) itself.

In a formula parallel to (4), we define the (symmetric) associated metric tensor g^{rs} through

$$dx^r = g^{rs}dx_s \ , \tag{7}$$

valid again in conjunction with an arbitrary vector in any coordinates. Using this and (5), we obtain an expression parallel to (2):

$$ds^2 = g^{rs}dx_r dx_s = \bar{g}^{ij}d\bar{x}_i d\bar{x}_j \ . \tag{8}$$

One can now proceed in a similar fashion that led to (3), except that (6b) should take the place of (1b). The result is

$$g^{rs} = (\partial x^r/\partial\bar{x}^i)(\partial x^s/\partial\bar{x}^j)\bar{g}^{ij} \ , \tag{9}$$

which also represents the transformation for a general second-order

contravariant tensor h^{rs}. Upon comparing it with (1a), one confirms that the contravariant tensor components transform like the contravariant vector components for each index.

One could form an invariant using a general second-order mixed tensor in conjunction with two independent vectors dx, dy as in

$$h_s^r dx_r dy^s = \overline{h}_j^i d\overline{x}_i d\overline{y}^j \ ,$$

which leads, with the use of both (1b) and (6b), to

$$h_s^r = (\partial x^r / \partial \overline{x}^i)(\partial \overline{x}^j / \partial x^s)\overline{h}_j^i \ . \tag{10}$$

This quantity transforms in the manner of contravariant vector components for the upper index and covariant vector components for the lower index. If "h" is replaced by "δ", the Kronecker delta is confirmed to be a mixed tensor.

According to §1 on page 9 in [2], a matrix can be regarded as a tensor if it transforms in the manner of vector components for each index. Since the above transformations for the tensors "g" or "h" can be proven just as easily for any number of indices, and since the space dimensions are now arbitrary rather than three, the correspondences involving (3×3) matrices could be generalized to correspondences involving higher-dimensional arrays. Similar generalizations could be made with regard to other tensor quantities. For example, $\partial x^r / \partial u^\alpha$ no longer corresponds merely to a (3×2) matrix as would be the case in the context of [2]. Since these as well as more complicated expressions having a higher number of indices may in turn be contracted

-26-

with various tensors, an analogy with higher-dimensional array operations could be established. The purpose of the foregoing has been to indicate that tensor concepts can be greatly expanded and have potentially few limitations. Although the present study is limited to second-order tensors at the most, these tensors can be related to matrices of general dimensions such as ($n \times n$), ($n \times u$), ($u \times u$), etc., instead of merely (3×3), (3×2) and (2×2). Since the tensors encountered herein will be expressed in terms of ortho-normal vectors in space and/or surface coordinates, the interpretation and treatment of the corresponding matrices will be particularly simple.

Due to the property of lowering or raising indices by the metric tensor or the associated metric tensor, respectively, one can write for arbitrary vectors A, B:

$$A_i B^i \equiv \delta_i^p A_p B^i = A^j B_j \equiv g^{pj} A_p g_{ji} B^i \quad ,$$

and thus

$$(\delta_i^p - g^{pj} g_{ji}) A_p B^i = 0 \quad .$$

Since A and B are independent, $A_p B^i$ represents n^2 independent coefficients connecting the n^2 components of the expression within the parentheses (it is inconsequential whether some or all of the components are distinct). It then follows that this expression must be zero for each component, that is,

$$g^{pj} g_{ji} = \delta_i^p \quad . \tag{11}$$

That this and other tensor equations are valid in any coordinates can be readily confirmed upon the multiplication by the appropriate transformation factors.

The orthonormal vectors ℓ, j, ... , ν, ... have the well-known property,

$$\ell_r \ell^r = 1 \ , \qquad \ell_r j^r = 0 \ , \ ... \ , \qquad \ell_r \nu^r = 0 \ , \ ... \ , \qquad (12)$$

which stems from the definition of the scalar product in n dimensions in analogy to its three-dimensional counterpart. In using these invariants, an arbitrary space vector A can always be described through

$$A^r = a\ell^r + bj^r + ... + k\nu^r + ... \ , \qquad (13a)$$

$$a = \ell_r A^r \ , \quad b = j_r A^r \ , \ ... \ , \quad k = \nu_r A^r \ , \ ... \ , \qquad (13b)$$

which enables us to write

$$g_{rs} A^r A^s \equiv A_s A^s = (\ell_r \ell_s + j_r j_s + ... + \nu_r \nu_s + ...) A^r A^s \ . \qquad (13c)$$

When subtracting the right-hand side from the left-hand side in (13c), we obtain a symmetric expression contracted by $A^r A^s$ which should result in zero (an invariant). Similar to the cases treated earlier we conclude that this symmetric expression must be zero and, therefore, that

$$g_{rs} = \ell_r \ell_s + j_r j_s + ... + \nu_r \nu_s + ... \qquad (14)$$

Due to the above symmetry, the double contraction by the arbitrary vector A in (13c) is sufficient without the necessity of bringing a second arbitrary vector into the picture.

Upon interchanging the upper and lower indices in (13a-c), similarly we obtain

$$g^{rs} = \ell^r \ell^s + j^r j^s + \ldots + \nu^r \nu^s + \ldots \tag{15}$$

With the aid of (14) and (15), equation (11) yields

$$\delta^r_s = \ell^r \ell_s + j^r j_s + \ldots + \nu^r \nu_s + \ldots , \tag{16}$$

where use has been made of the orthonormal properties (12). If (1a) is applied to all of the orthonormal vectors, one can form

$$\ell^r \bar{\ell}_t + j^r \bar{j}_t + \ldots + \nu^r \bar{\nu}_t + \ldots = (\partial x^r / \partial \bar{x}^s)(\bar{\ell}^s \bar{\ell}_t + \bar{j}^s \bar{j}_t + \ldots + \bar{\nu}^s \bar{\nu}_t + \ldots).$$

Upon taking advantage of (16) in $\bar{x}^r$ coordinates, this is written as

$$\partial x^r / \partial \bar{x}^t = \ell^r \bar{\ell}_t + j^r \bar{j}_t + \ldots + \nu^r \bar{\nu}_t + \ldots \tag{17}$$

The reversed partial derivatives would follow from interchanging the overbars in (17).

2.2 Relations Involving Extrinsic Properties of Surfaces

We first rewrite some of the formulas derived in the preceding section which are immediately applicable in a u-dimensional space viewed here as a surface. In particular, (4), (7), (12), (14), (15) and (16) together with (11) are transcribed, respectively, as

$$du_\beta = a_{\beta\alpha} du^\alpha \ , \qquad du^\alpha = a^{\alpha\beta} du_\beta \ , \tag{18a,b}$$

$$\ell_\alpha \ell^\alpha = 1 \ , \qquad \ell_\alpha j^\alpha = 0 \ , \ \dots \ , \tag{19}$$

$$a_{\alpha\beta} = \ell_\alpha \ell_\beta + j_\alpha j_\beta + \dots \ , \qquad a^{\alpha\beta} = \ell^\alpha \ell^\beta + j^\alpha j^\beta + \dots \ , \tag{20a,b}$$

$$\delta^\alpha_\beta = a^{\alpha\gamma} a_{\gamma\beta} = \ell^\alpha \ell_\beta + j^\alpha j_\beta + \dots \ , \tag{21}$$

where use is made of the orthonormal vectors ℓ, j, etc., written now in terms of surface coordinates.

The same physical vector of length ds' lying in the surface can be represented by space as well as by surface components, both contravariant and covariant (one kind can be obtained from the other via the appropriate metric tensor), as follows:

$$dx'^r = a\ell^r + bj^r + \dots \ , \qquad dx'_s = a\ell_s + bj_s + \dots \ , \tag{22a,b}$$

$$du^\alpha = a\ell^\alpha + bj^\alpha + \dots \ , \qquad du_\beta = a\ell_\beta + bj_\beta + \dots \ ; \tag{23a,b}$$

from here the scalar invariants a, b, etc., can be found through the self-evident contractions of that vector with ℓ, j, etc., respectively. The invariant ds'^2 is given in several equivalent ways in the expression

-30-

$$ds'^2 = dx'^r dx'_r = g_{rs} dx'^r dx'^s = g^{rs} dx'_r dx'_s$$

$$= du^\alpha du_\alpha = a_{\alpha\beta} du^\alpha du^\beta = a^{\alpha\beta} du_\alpha du_\beta = a^2 + b^2 + \dots \qquad (24)$$

This could represent the first step in connecting space and surface tensors, but we shall proceed along slightly different lines. One notices that the result of (24) necessarily agrees with the Cartesian formulation which merely introduces a special coordinate system for expressing ℓ, j, etc.

In Chapter 6 of [2], dealing with two-dimensional surfaces embedded in a three-dimensional space, several forms describing a surface are presented. The first, or Gauss' form, has the most general characteristics and is adopted for the present development. It now expresses each of the n space coordinates x^r as some function of the u surface coordinates u^α, namely

$$x^r = x^r(u^\alpha) ; \qquad r = 1,2,\dots,n ; \qquad \alpha = 1,2,\dots,u .$$

The ordinary rule of differentiation for this system of equations yields the following linear relation between the space components dx'^r and the surface components du^α of a vector lying in the surface:

$$dx'^r = (\partial x^r / \partial u^\alpha) du^\alpha . \qquad (25)$$

This is essentially equation 6.01 of [2] adapted to the present context, and it is a counterpart of (1a) where the vector dx is unrestricted.

Since dx' is the same physical vector as du expressed in different coordinates, the above is one of the basic formulas relating the space and the surface components of such vectors. If we apply this formula to ℓ, j, etc., in analogy to the steps that have lead to (17) we obtain

-31-

$$\ell^r \ell_\beta + j^r j_\beta + \ldots = \partial x^r / \partial u^\beta ,$$

where (21) has been taken into account for the right-hand side. This vectorized form closely resembles the tensor equation 6.09 in [2] where, however, the derivation was based on a relation of the type (24). Clearly, it transforms like a space tensor in the contravariant indices and like a surface tensor in the covariant indices. Upon multiplying it (with the index r changed to s) by dx'_s as given in (22b) and considering (23b), one finds

$$du_\beta = (\partial x^s / \partial u^\beta) dx'_s , \tag{26}$$

which resembles 6.08 in [2] derived again differently.

Introducing the notation A^r_α for the partial derivatives $\partial x^r / \partial u^\alpha$, we have

$$A^r_\alpha \equiv \partial x^r / \partial u^\alpha = \ell^r \ell_\alpha + j^r j_\alpha + \ldots , \tag{27}$$

whose components could be found explicitly from the Gauss' form of the surface. Using this notation we write (25) and (26) as

$$dx'^r = A^r_\alpha du^\alpha , \qquad du_\beta = A^s_\beta dx'_s , \tag{28a,b}$$

which is confirmed also with the relations (22a)-(23b) together with (27). Similar formulas can be written in terms of the vectors ℓ, j, etc., namely

$$\ell^r = A^r_\alpha \ell^\alpha , \qquad j^r = A^r_\alpha j^\alpha , \ldots$$

In considering an arbitrary vector dx'' from the subspace orthogonal to the surface, it follows that

$$A^S_\beta dx''_S = 0 .$$

(29)

Next we define a different vectorized expression,

$$Q^\alpha_r = \ell^\alpha \ell_r + j^\alpha j_r + \dots ,$$

(30)

which transforms like a surface tensor in the contravariant indices and like a space tensor in the covariant indices. Although "Q" may not be given directly, it is linked to "A" in a simple relationship to be shown later in this section. Upon changing the indices α, r to β, s in (30), multiplying this equation by (23b) and considering (22b), one has (31a) below, while the consideration of (30) and (22a), (23a) leads to (31b):

$$dx'_s = Q^\beta_s du_\beta , \qquad du^\alpha = Q^\alpha_r dx'^r .$$

(31a,b)

With regard to the vector dx", it now holds that

$$Q^\alpha_r dx''^r = 0 .$$

(32)

In considering the importance of the model surface in the upcoming development, it is useful to introduce a symmetric tensor g'_{rs} having the role of the space metric tensor, but only for the vectors restricted to this surface:

$$g'_{rs} = \ell_r \ell_s + j_r j_s + \dots$$

(33a)

Multiplying (22a) by this tensor yields (22b). Thus we see that (4) is obtained with all three tensors primed and that the primes can similarly be added to (2) and (5), etc., which confirms the desired characteristics of (33a). We note that since g_{rs} lowers the indices in conjunction with

any space vectors, the use of this tensor is sufficient -- but not necessary -- for lowering the indices in conjunction with the vectors lying in the surface. The tensor of the lowest rank to accomplish this is g'_{rs}, which can thus be said to be the "necessary metric tensor" in this context.

Similar statements with self-evident modifications can be also made with regard to the tensor g'^{rs}, namely

$$g'^{rs} = \ell^r \ell^s + j^r j^s + \ldots , \tag{33b}$$

which can be called the "necessary associated metric tensor". For example, when this tensor is applied to (22b), equation (7) is recovered with all three quantities primed, etc. It is worth mentioning that the orthonormal vectors ℓ, j, etc., are not given, nor is it necessary to express them in any way, but they are used as geometrical tools in unearthing the relationships between various tensor quantities and formulating them in tensor equations valid in any coordinates.

In using (27) and (30) together with the vectorized forms for the (associated) metric tensors already encountered, the following identities are readily established:

$$A^r_\alpha = g'^{rs} Q^\beta_s a_{\beta\alpha} = g^{rs} Q^\beta_s a_{\beta\alpha} , \tag{34}$$

$$Q^\alpha_r = a^{\alpha\beta} A^s_\beta g'_{sr} = a^{\alpha\beta} A^s_\beta g_{sr} . \tag{35}$$

As well, due to (21) the identity below holds true:

$$Q^\alpha_r A^r_\beta = \delta^\alpha_\beta . \tag{36}$$

Recalling (33a,b) one also derives

$$A_\alpha^r Q_t^\alpha \equiv \ell^r \ell_t + j^r j_t + \ldots = g'^{rs} g'_{st} . \tag{37}$$

The final (non-vectorized) formulas giving the metric tensor and the
associated metric tensor in surface coordinates as well as the final
formulas giving the "necessary" counterparts of these tensors in space
coordinates will be presented in the next chapter via a geometrical com-
mutative diagram. The subsequent chapter will utilize them in establishing
the link between certain tensor transformations and the law of the
variance-covariance propagation.

3. GEOMETRICAL CONFIGURATION REFLECTING THE LEAST-SQUARES SETUP

In Section 2.2 only one surface has been considered, called the model surface, to which all of the formulas (18a)-(37) have been related in some way. Only equations (29) and (32) have featured a vector (dx") independent of this surface, in particular, an arbitrary vector from the subspace orthogonal to the surface at the point P. Such a vector has not yet been related to dx or to another surface. But if the vector dx is decomposed into dx' and $d\tilde{x}$, where the first lies in the model surface and the second lies in some other surface, the formulas from Section 2.2 can be adapted for $d\tilde{x}$ and the related quantities. Clearly, the "other surface" would have to be specified. The easiest configuration to handle occurs when this surface is identified with the subspace orthogonal to the model surface at P and becomes thus the "second surface" mentioned in Chapter 1. The metric and the associated metric tensors in surface coordinates can then be expressed in a vectorized form similar to (20a,b), where the orthonormal vectors ℓ, j, etc., are replaced by ν, etc. Just as the former, the latter also coincide with the physical vectors used earlier in the expressions of space relations.

According to the above stipulation, dx is decomposed into two orthogonal vectors. One, again denoted dx', lies in the model surface and the other, denoted dx", lies in the second surface. This brings about a number of advantages. For example, the relations (29) and (32) with this particular vector dx" -- as opposed to some quite arbitrary vector dx" as implied therein -- can be used to express dx' in surface coordinates

-36-

directly from dx, etc. In using the counterparts of the formulas in
Section 2.2 with regard to the second surface, dx" can likewise be expres-
sed in surface coordinates directly from dx. All the relationships stemming
from the present geometrical configuration will be seen as exceedingly
simple, especially if they should be compared with a procedure based on a
different second surface. But this configuration with all its advantages
and clearcut characteristics would be of little use in the adjustment
theory if it did not faithfully reflect the L.S. setup.

In order to prepare the ground for relating differential geometry to
the L.S. adjustment, dx is regarded as an observational vector. In the
absence of any errors, the observational vector would coincide with a "true"
vector restricted by the nature of the problem to what is called here the
model surface. In Fig. 1 showing a simplified situation in a three-
dimensional space (n=3, u=2), this "true" vector could be symbolized by an
arrow from P to Q^t. But since the "true" vector can never be known, it is
estimated by another vector which likewise must lie in the model surface
(this restriction is expressed by the pertinent mathematical model con-
sidered known). One such vector, called the model vector, could be
symbolized in Fig. 1 by an arrow from P to $\tilde{Q}$. An arrow from $\tilde{Q}$ to Q
(parallel-transported to P) would then represent estimates of errors in
the observational vector and would be called the error vector. The
observational vector can thus be said to be decomposed into a model vector
and an error vector, all belonging to the point P.

Within a small neighborhood of P considered, the model surface in

-37-

Fig. 1 is approximated by a plane. This figure then corresponds to Fig. 1.3.1 of [13], except that the observational vector (y) therein is associated with the metric in Cartesian coordinates whereas the vector dx above can be associated with a more general metric. But in other respects the present geometrical representation of the L.S. setup, including the notion of minimum length for the error vector as explained in the next paragraph, is consistent with Section 1.3 of [13]. One may notice that what is called here the second surface corresponds, according to Definition 16 on page 384 of the same reference, to an orthocomplement of the model surface at P in the underlying space.
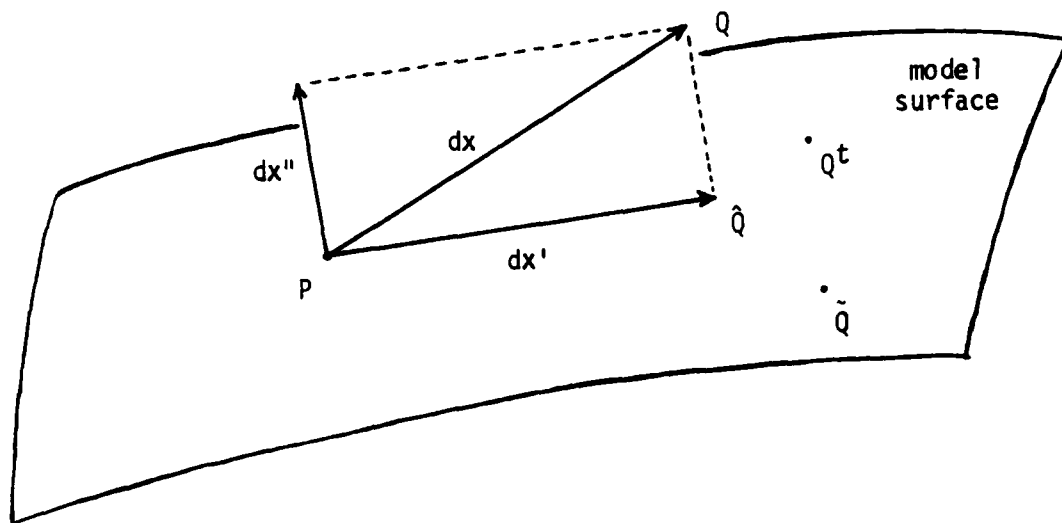
Figure 1

Symbolic representation of the observational vector dx,
model vector dx' and error vector dx"

Because the model surface contains the all-important model vector, the need for expressing certain extrinsic properties of this surface is easily understood. However, the usefulness of introducing another surface and its extrinsic properties into the derivations depends on the type of adjustment one intends to apply to the observational vector. If one stipulates that the length of the error vector should be a minimum, it follows that this vector must be orthogonal to the model surface. It thus lies in the second surface and is identified by the symbol dx" in agreement with the second paragraph of this chapter. Under this stipulation the decomposition of the observational vector is characterized by

$$dx_s = dx'_s + dx''_s , \qquad dx^r = dx'^r + dx''^r , \qquad (38a,b)$$

where the model vector dx' is the orthogonal projection of the observational vector dx on the model surface and the vector dx" is the orthogonal projection of dx on the second surface. This type of adjustment was anticipated in Fig. 1 (hence the point $\hat{Q}$), where the second "surface" is represented by the straight line through P perpendicular to the (two-dimensional) model surface. The minimum condition for the geometrical quantity ds''^2,

$$ds''^2 = dx''^s g_{sr} dx''^r = dx''_r g^{rs} dx''_s , \qquad (39)$$

corresponds to the minimum L.S. criterion for $V^T PV$ in either of the two identifications mentioned in Chapter 1. In particular, the first form in (39) represents the traditional identification and the second form represents the new identification.

The remarkable consequence of the minimum property (39), and thus of the introduction of the second surface and the decomposition (38a,b), is that both the model vector and the error vector fulfill the same relationships, except that two different surfaces -- and the corresponding two sets of tensor quantities -- are present. Once all the desired relationships involving the model surface are derived, their counterparts involving the second surface can readily be transcribed using simple changes in notations. The final formulas do not contain the vectorized expressions whose main role has been to facilitate the derivations. Therefore, such expressions need not be listed for the second surface, except for the following four which can be used to derive several "cross" relations involving both surfaces:

$$S_L^r \equiv \partial x^r / \partial u^L = \nu^r \nu_L + \ldots , \tag{40}$$

$$B_r^L = \nu^L \nu_r + \ldots ; \tag{41}$$

$$g''_{rs} = \nu_r \nu_s + \ldots , \qquad g''^{rs} = \nu^r \nu^s + \ldots \tag{42a,b}$$

Equations (40) and (41) are the counterparts of (27) and (30), respectively, and the "necessary" tensors in (42a,b) are the counterparts of (33a,b). The tensors in (42a,b) are related to dx" in the same way as the tensors in (33a,b) have been related to dx'. Upon recalling the vectorized forms of all the pertinent (associated) metric tensors, parallel to (38a,b) one can write

$$g_{sr} = g'_{sr} + g''_{sr} , \qquad g^{rs} = g'^{rs} + g''^{rs} . \tag{43a,b}$$

As an immediate consequence of the decomposition in (38a,b), we have

$$du_\beta = A_\beta^S dx_S' = A_\beta^S dx_S \ , \qquad du^\alpha = Q_r^\alpha dx'^r = Q_r^\alpha dx^r \ , \qquad (44a,b)$$

where (28b) and (29) have been used in (44a), and (31b) with (32) have been used in (44b). Similar formulas could be written for du_M and du^L involving the second surface. A number of other relations, including those for the (associated) metric tensors, could be derived from the vectorized expressions in conjunction with either surface. But such an exercise is avoided. Instead, a geometrical commutative diagram is presented in Fig. 2 from which all the final relationships in terms of first- and second-order tensors can be read. Some have already been derived and all the others can easily be verified by means of the usual vectorized expressions. The diagram offers a clearcut illustration of the complete analogy between the formulas pertaining to one or the other surface.

A detailed description of Fig. 2 is in order. First of all, the vector components (marked in boxes) also represent the pertinent spaces or subspaces. Thus the boxes representing dx' and dx" should be imagined as completely filling the box representing dx in both the contravariant and covariant versions. The second-order tensors acting as linear transformation operators are designated by arrows. The heavier lines identify the quantities which, in the corresponding L.S. setup, are considered known a priori. The dashed arrows identify the tensors regarded as intermediate, whose presence may or may not be desired in the final results (if it is not desired, they can be expressed by the appropriate contractions of other tensors). But this diagram could be imagined just as well with all the lines drawn solid

-41-

and of the same thickness. It would then illustrate even better the analogy between the relationships involving the model surface and the second surface.

The arrows with dots in Fig. 2 can be used in two ways, i.e., the dots can either be considered as an integral part of the arrow (hence "complete arrow") or be disregarded (hence "shortened arrow"). The heavy vertical arrows g_{sr} and g^{rs} are depicted as connecting the boxes representing dx. But since these boxes encompass both dx' and dx", the heavy vertical arrows can be used in conjunction with any of dx, dx' or dx". However, one has to make sure that both ends of these arrows pertain to the same vector (this is not necessarily the case with the arrow g'^{rs}, for example, whose tip corresponds to dx' but whose base may correspond to dx as one of the two choices). Thus only the complete and not the shortened g'- or g"-arrows can be substituted for by the corresponding g-arrows.

When working with the present diagram, one proceeds in general against the direction of the arrow(s) when expressing one quantity in terms of some other(s). When considering dx_s , for example, we notice that the tip of only one arrow touches its box. Accordingly, it can be connected to another quantity only through the arrow g_{sr} whose base touches dx^r. By proceeding against the direction of this arrow dx_s is expressed as in (4). In seeking to express dx^r one similarly obtains (7). In the same vein, the diagram gives du_β in terms of du^α, and du^α in terms of du_β, as in (18a) and (18b), respectively. The vector components du^L and du_M are expressed in complete analogy to these equations, except that the left-hand side of the diagram, rather than its right-hand side, is utilized. Due to such

-42-

analogies, we can concentrate on the right-hand side of the diagram when explaining its functioning.

We notice that there are different ways available in formulating certain vector components in terms of others. For example, du^α can also be written with the aid of the arrow Q_r^α as in (44b), where the first and second equalities correspond to the complete and shortened arrows, respectively. The vector components dx'^r can in turn be expressed by means of the shortened arrow g'^{rs} in conjunction with dx_s, or by means of the complete arrow g'^{rs} in conjunction with dx'_s resulting in equation (7) with all the quantities primed as was already mentioned following (33b). In the latter case the complete g'-arrow can be replaced by the corresponding g-arrow. Accordingly, du^α can be written in terms of dx'_s through the contraction of Q_r^α and g'^{rs} or, equally well, through the contraction of Q_r^α and g^{rs}. Exploiting only some of the above possibilities, we write

$$dx'^r = g'^{rs}dx'_s = g'^{rs}dx_s , \qquad du^\alpha = Q_r^\alpha g'^{rs}dx_s , \qquad (45a,b)$$

as two examples of a great number of transformation formulas that can be obtained with the aid of Fig. 2.

In assessing the generality of the present geometrical diagram, we should be able to recover the correct relationships also when selecting some very special paths. For example, one can choose to return to the same box. In these and similar operations, the indices may be renamed but must always be (re)arranged in such a way that the free indices do not repeat themselves and the dummy indices repeat themselves only once. Thus, if we strive to express du^α in terms of itself by following the two shortest

-43-

paths (first path: down and then up, second path: left and then right), or express dx'^r in terms of itself, with the appropriate index changes we have

$$du^\alpha = a^{\alpha\beta}a_{\beta\gamma}du^\gamma = Q^\alpha_r A^r_\delta du^\delta \ , \qquad dx'^r = A^r_\alpha Q^\alpha_s dx'^s \ , \qquad (46a,b)$$

with (46a) verified by (21) and (36), respectively. In following the two longest paths (without repeating any parts thereof) as another example involving du^α , we can write

$$du^\alpha = (Q^\alpha_r g'^{rs}Q^\beta_s)(A^p_\beta g'_{pq}A^q_\gamma)du^\gamma = (a^{\alpha\beta}A^s_\beta g'_{sr})(g'^{rp}Q^\gamma_p a_{\gamma\delta})du^\delta \ , \qquad (46c)$$

where any or all of the g'-arrows can be replaced by their unprimed counterparts. We note that the expressions in parentheses in (46c) are equal to their counterparts in (46a), as can be readily asserted from the vectorized forms. In fact, for the last terms in (46a,c) these equivalences have already been established in (34) and (35). Two paths of intermediate length could also be considered, one describing the quad counterclockwise (starting and ending at du^α) and the other describing it clockwise. Other paths are possible, such as proceeding half way around the quad and returning to du^α. One could also adopt any of the paths just mentioned and lengthen it by repeating some or all of the arrows. But such repeating arrows are of little use in general.

The second-order tensors in the diagram can be expressed in terms of others by starting at the tip of the arrow whose formulation is sought and proceeding against any chosen arrows until the base of the original arrow is reached. A simple application of this rule yields the indicated equalities

-44-

between (46a) and (46c) as discussed below the latter. When applied to g'_{sr} and g'^{rs}, this rule gives the formulas which are even simpler, in that only one set of equalities exists in either case.

In a somewhat more complicated application, an arrow could be expressed in terms of itself. In most cases of this kind at least one arrow would have to be repeated. For example, we have

$$(47a,b,c,d)$$

$$Q^\alpha_r = (Q^\alpha_p A^p_\delta) Q^\delta_r \ , \qquad a^{\alpha\beta} = (a^{\alpha\gamma} a_{\gamma\delta}) a^{\delta\beta}, \qquad A^r_\alpha = A^r_\beta Q^\beta_s A^s_\alpha, \qquad g'^{rs} = g'^{rp} g'_{pq} g'^{qs}.$$

In (47d), g'^{rp} could be replaced by g^{rp}, for example, so that no arrows on the right-hand side would repeat themselves, but the g'-arrow would still be written in terms of itself (see g'^{qs} on the right-hand side). If also g'^{qs} were replaced by g^{qs} the g'-arrow would not be written in terms of itself, but the g-arrow would be repeated. If we replaced all three g'-arrows on the right-hand side by their unprimed counterparts, we would have to similarly replace the g'-arrow on the left-hand side because of the following simple rule: when no other arrow except the g-arrows are involved in an expression, they can only represent the heavy vertical arrows of Fig. 2. The examples treated in this paragraph have little use in practice, but they have been presented in order to demonstrate that the diagram can handle any combinations of arrows, including even their multiple repetitions.

Another class of formulas can be obtained upon expressing a combination of arrows forming one "long" arrow. Proceeding even here from the tip toward the base we have, as two examples,

$$g'^{rs} Q^\beta_s = A^r_\alpha a^{\alpha\beta} \ , \qquad Q^\beta_s a_{\beta\alpha} = g'_{sr} A^r_\alpha \ ,$$
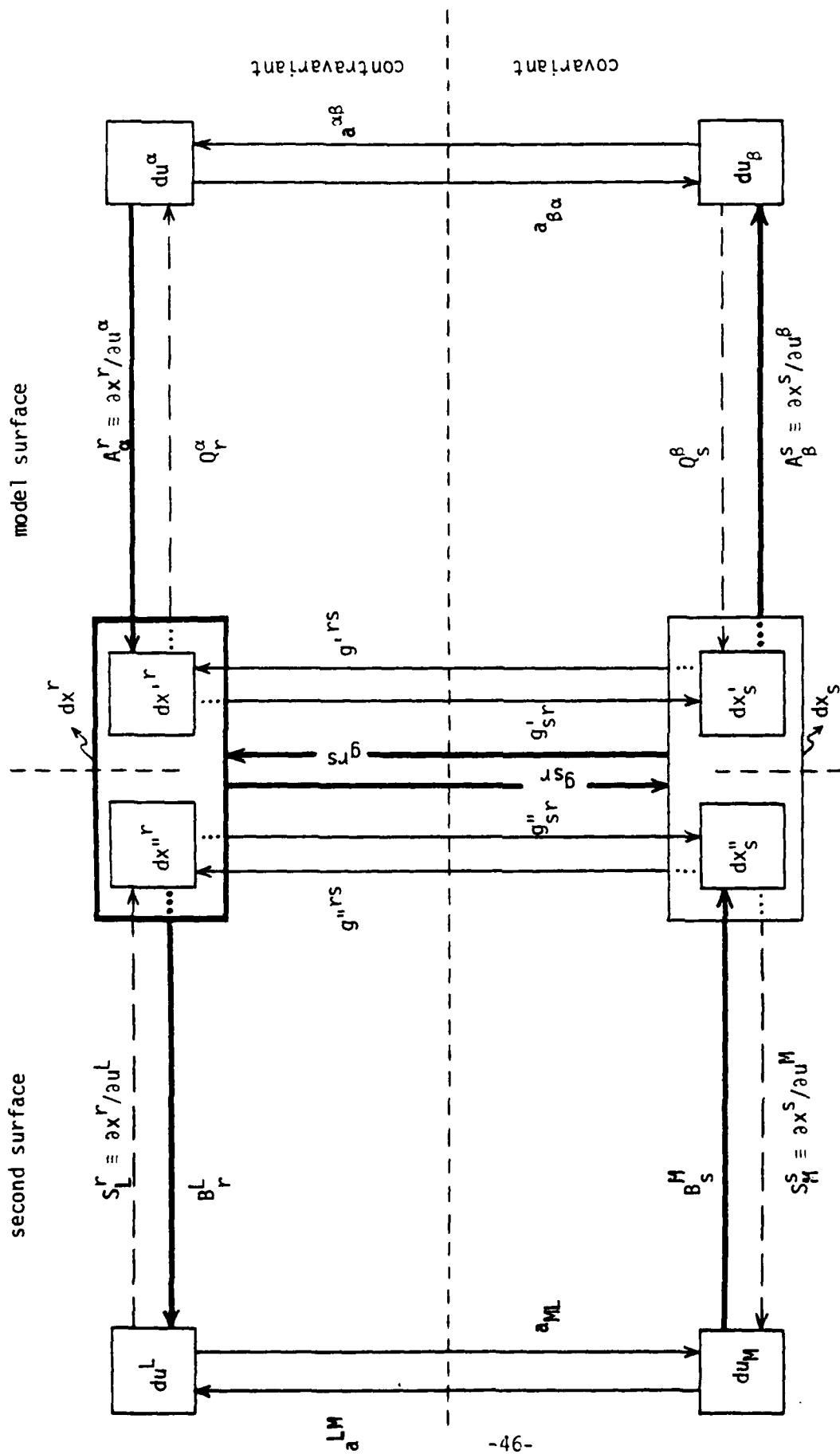
-45-

Figure 2

Geometrical commutative diagram

where the g'-arrows can again be replaced by their unprimed counterparts.
Some useful relationships can be arrived at through special combinations
of arrows reaching the tip of the initial arrow. An example can be found
in (47a,b), where the expressions in parentheses are the same tensor.
According to (21) this tensor is δ^{α}_{δ}, which then confirms (36). Another
example of this kind is equation (37). More complicated examples of these
special combinations can be found, such as using more than one arrow in
one sense and then returning to the tip in the opposite sense, etc.
Although such elaborate cases will not be needed in practice, it is com-
forting to verify that the diagram can handle them as well.

As the next step, we consider both surfaces when carrying out operations
which could be called "cross-contractions". It can be readily verified that
the contraction of a tensor (either order) associated with one surface by a
tensor associated with the other surface yields a zero tensor. In terms of
the diagram, when an arrow is followed by a box or an arrow across the
dashed vertical line, the result is zero. The heavy vertical lines are not
a part of such cross-contractions since they are considered to be associated
with both surfaces. Two examples of the cross-contractions are equations
(29) and (32). The counterparts of these examples follow when proceeding
from the second surface to the model surface. As other examples, we have

$$g'^{rs}dx''_s = 0, \quad g'_{sr}dx''^r = 0 \ldots dx'^S g'_{sr}dx''^r \equiv dx'^S g_{sr}dx''^r \equiv dx'_r dx''^r = 0,$$

together with their counterparts obtained again by proceeding in the other
sense. As a consequence of the cross-contractions, a tensor belonging to
the model surface contracted with the components of dx' is equal to the

-47-

same tensor contracted with the components of dx, as seen e.g. in (44a,b) or (45a). This feature is embodied in the properties of the complete and shortened arrows explained earlier.

With regard to the cross-contractions involving the arrows alone, we can write a number of relations such as the selected six below:

$$B_r^L A_\alpha^r = 0 \ , \qquad Q_r^\alpha S_L^r = 0 \ , \qquad Q_r^\alpha g'^{rs} B_s^M \equiv Q_r^\alpha g^{rs} B_s^M = 0 \ ; \qquad (48a,b,c)$$

$$g''_{sr} A_\alpha^r = 0 \ , \qquad B_r^L g'^{rs} = 0 \ , \qquad g'_{sr} g''^{rt} = 0 \ . \qquad (48d,e,f)$$

Similar to the preceding paragraph we can state that as a consequence of the cross-contraction properties, a contraction of a tensor belonging to the model surface with a complete g'-arrow is equal to the contraction of the same tensor with the corresponding g-arrow. But the replacement of a complete g'-arrow by its unprimed counterpart has also been explained earlier. Clearly, the above rules hold just as well when the second surface is considered in lieu of the model surface.

Due to the additive relations (38a,b) and (43a,b), various contractions among the first- and second-order primed tensors can be added algebraically to their double-primed counterparts, the result being the corresponding expression without primes. The formulas (38a,b) themselves are the simplest examples with regard to the vector components alone (without any contractions), and the formulas (43a,b) themselves are the simplest examples with regard to the second-order tensors alone. This rule is particularly easy to follow in the diagram. For tensor-vector contractions we have

$$g'_{sr} dx'^r + g''_{sr} dx''^r = g_{sr} dx^r \ , \qquad g'^{rs} dx'_s + g''^{rs} dx''_s = g^{rs} dx_s \ , \qquad (49a,b)$$

-48-

where the g'- and g"-arrows could be replaced by their unprimed counter-parts as usual. These equations are essentially (38a,b). And tensor-tensor contractions result in

$$g'^{rs}g'_{st} + g''^{rs}g''_{st} = A^r_\alpha Q^\alpha_t + S^r_L B^L_t = g^{rs}g_{st} = \delta^r_t .$$ (50)

The above rule applies between the first and the third entity in (50). In the first entity one of the g'-arrows and one of the g"-arrows could be replaced by the appropriate g-arrows, but clearly not both (this also transpires from the discussion that followed 47d). The first equality in (50) is the result of a simple return to the tip performed for either term, and the last equality in (50) is equation (11). Nothing new would be gained in an expression resulting from the path "up and then down" as compared to the path "down and then up" chosen in (50).

Finally, an expression for ds''^2 is developed in conjunction with either surface. Upon considering the model surface, the geometrical configuration yields

$$ds''^2 = ds^2 - ds'^2 ,$$

from which it follows that

$$ds''^2 = dx^s g_{sr} dx^r - du^\beta du_\beta , \qquad ds''^2 = dx_r g^{rs} dx_s - du_\alpha du^\alpha .$$ (51a,b)

The first terms on the right-hand sides of (51a,b) are essentially (2) and (8), respectively, while the second terms follow from (24). Expressing du_β in (51a) and du^α in (51b) from the diagram, we have

-49-

$$ds''^2 = dx^s g_{sr} dx^r - du^\beta A^s_\beta g_{sr} dx^r , \quad ds''^2 = dx_r g^{rs} dx_s - du_\alpha Q^\alpha_r g^{rs} dx_s .$$

In the second terms on the right-hand sides of (52a,b), the g-arrows could have been written as g'-arrows and/or the dx vectors could have been written as the dx' vectors. With regard to the second surface, a counterpart of (24) yields

$$ds''^2 = du^M a_{ML} du^L , \qquad ds''^2 = du_L a^{LM} du_M . \tag{53a,b}$$

4. TENSOR STRUCTURE AND THE LAW
OF VARIANCE-COVARIANCE PROPAGATION

We have seen in Chapter 1 that the traditional identification represented by [11] has been limited to linking the variance-covariance propagation law to the formula for the associated metric tensor based on a coordinate transformation in one and the same n-dimensional space. This concept will now be generalized and extended. First, we confirm the propagation law in the above situation called n-n, representing the transformation within the same space from n coordinates x^r to n coordinates $\bar{x}^r$. Next, the situation n-u is addressed, corresponding to the transformation from n coordinates x^r to u coordinates u^α. In a subsequent step we treat the situation u-n, corresponding to the transformation from u coordinates u^α to n coordinates x^r. The last situation treated is u-ñ, corresponding to the transformation from u coordinates u^α to $\tilde{n}$ coordinates $\tilde{x}^\Omega$; the upper-case Greek letters are introduced temporarily to distinguish the indices in an ñ-dimensional space from those in the other spaces already encountered. The situation u-u need not be treated because it is the same as n-n, except that the one space considered is u- rather than n-dimensional. As well, nothing new would ensue from replacing u dimensions by r, etc.

In order to supplement the links between the C-matrices and the associated metric tensors, the P-matrices in all four situations above will similarly be linked to the metric tensors. The resulting relationships will unearth the fact that one can formulate the "law of weight propagation" in a complete analogy to the well-known law of variance-covariance propagation. The above statements apply to the traditional identification. In a

subsequent derivation the new identification will replace the traditional one in all four situations. The outcome will reveal that with proper substitutions of second-order tensors and tensor-related quantities such as $\partial \bar{x}^i / \partial x^r$ by matrices, exactly the same formulas for both C- and P-matrices are recovered in either identification.

It should be pointed out that although in most cases the notations introduced for vectors and matrices are similar to the notations used later in describing the parametric method, the subjects of the variance-covariance propagation and the "weight propagation" are presented independently of any L.S. considerations. Indeed, it is not important at this point whether or not the decomposition of dx into two orthogonal vectors dx' and dx" corresponds to the L.S. adjustment. The important property we wish to expose is that for every encountered relationship between the vector components there exists a relationship between the corresponding (associated) metric tensors which fits perfectly the variance-covariance propagation law. This in itself could constitute motivation for using differential geometry with tensor structure in an analysis of various L.S. methods and their properties.

4.1 Metric Relations Corresponding to Various Types of Vector Relations

The vector relations depicting the first three of the four situations described above are, in the same order, (1b), (6b), (44b), (44a), (28a) and (31a). All except the first two can be read from the diagram of Fig. 2. With regard to the situation n-n, we introduce the notations $D_r^i \equiv \partial \bar{x}^i / \partial x^r$ and $E_m^s \equiv \partial x^s / \partial \bar{x}^m$, and transcribe the formulas giving the (associated) metric tensors in (9) and (3), respectively, with the overbars interchanged. The vector formulations below are separated from their (associated) metric tensor counterparts by a few dots. Although these three situations require no more than a simple transcription, the fourth situation will necessitate a separate derivation.

Situation n-n.

$$d\bar{x}^i = D_r^i dx^r \quad \ldots\ldots \quad \bar{g}^{ij} = D_r^i g^{rs} D_s^j \;, \tag{54a,b}$$

$$d\bar{x}_m = E_m^s dx_s \quad \ldots\ldots \quad \bar{g}_{mk} = E_m^s g_{sr} E_k^r \;. \tag{55a,b}$$

Situation n-u.

$$du^\alpha = Q_r^\alpha dx'^r = Q_r^\alpha dx^r \quad \ldots\ldots \quad a^{\alpha\beta} = Q_r^\alpha g'^{rs} Q_s^\beta = Q_r^\alpha g^{rs} Q_s^\beta, \tag{56a,b}$$

$$du_\beta = A_\beta^s dx'_s = A_\beta^s dx_s \quad \ldots\ldots \quad a_{\beta\alpha} = A_\beta^s g'_{sr} A_\alpha^r = A_\beta^s g_{sr} A_\alpha^r. \tag{57a,b}$$

Situation u-n.

$$dx'^r = A_\alpha^r du^\alpha \quad \ldots\ldots \quad g'^{rs} = A_\alpha^r a^{\alpha\beta} A_\beta^s \;, \tag{58a,b}$$

$$dx'_s = Q_s^\beta du_\beta \quad \ldots\ldots \quad g'_{sr} = Q_s^\beta a_{\beta\alpha} Q_r^\alpha \;. \tag{59a,b}$$

<u>Situation u-ñ.</u> *Suppose now that* $\tilde{n}$ *numbers grouped as contravariant vector components* $d\tilde{x}'^{\Lambda}$ *are expressed via linear combinations of the components* du^{α}. *Such a relation can be written in a form similar to (58a) with another new notation,* $\tilde{A}^{\Lambda}_{\alpha}$. *We recall that* A^{r}_{α} *in (58a), in its matrix representation, is assumed to have the full (column) rank u. If the rank of the new quantity* $\tilde{A}^{\Lambda}_{\alpha}$ *is also u, the derivation can proceed without any further delay (see the next paragraph). If this rank is smaller, the indices* Λ *are extended as much as necessary through the addition of essentially arbitrary components, resulting in an extended quantity* $\tilde{A}^{\Lambda}_{\alpha}$ *and, accordingly, extended vector components* $d\tilde{x}'^{\Lambda}$. *In matrix terminology, arbitrary rows are being added until the rank of the extended matrix* $\tilde{A}$ *reaches u. The reasons for the original rank deficiency do not matter (the original number* $\tilde{n}$ *could have been smaller than u, as one example). In the final result the added (arbitrary) values will be discarded, but the full rank is needed if the interpretation and the derivation should proceed in analogy to (58a,b).*

At this point the interpretation of the new (and possibly extended) expression is similar to that of (58a), except that dx' and A are now attributed the symbol "~". In particular, $d\tilde{x}'$ is an $\tilde{n}$-dimensional vector lying in the surface and expressed in the space components, whereas du is the same vector expressed in the surface components. The nature of the $\tilde{n}$-dimensional space in which the surface is embedded is of little interest here. The important fact is that the derivation of the desired associated metric tensor can proceed along the same lines as the derivation of (58b). This leads to the relationships closely resembling (58a,b). Next we drop

the added dimensions and components, if any, and present the results as

$$d\tilde{x}'^{\Lambda} = \tilde{A}_{\alpha}^{\Lambda}du^{\alpha} \quad \ldots \ldots \quad \tilde{g}'^{\Lambda\Omega} = \tilde{A}_{\alpha}^{\Lambda}a^{\alpha\beta}\tilde{A}_{\beta}^{\Omega} \; . \tag{60a,b}$$

Clearly, the case where the components of $d\tilde{x}'$ are written as linear functions of the components of dx' is included in (60a,b) because the components of dx' are linear functions of the components of du as is shown in (58a). In such a case we would write

$$d\tilde{x}'^{\Lambda} = \tilde{H}_{r}^{\Lambda}dx'^{r} \quad \ldots \ldots \quad \tilde{g}'^{\Lambda\Omega} = \tilde{H}_{r}^{\Lambda}g'^{rs}\tilde{H}_{s}^{\Omega} \; . \tag{60'a,b}$$

According to (58a), equation (60a) is equivalent to (60'a) with

$$\tilde{A}_{\alpha}^{\Lambda} = \tilde{H}_{r}^{\Lambda}A_{\alpha}^{r} \; ,$$

which, together with (58b), shows that equation (60b) is equivalent to (60'b).

In analogy to the above development we can consider $\tilde{n}$ components $d\tilde{x}'_{\Omega}$ expressed in terms of du_{β} , similar to (59a) with another new notation, $\tilde{Q}_{\Omega}^{\beta}$. If the rank of the latter is smaller than u, an extension (this time in Ω) again takes place. The interpretation of the new (and, eventually, extended) expression is similar to that of (59a); the description in the preceding paragraph can be adopted here as well. Accordingly, the relationships parallel to (59a,b) are obtained where the added dimensions and components, if any, are dropped. The result reads

$$d\tilde{x}'_{\Omega} = \tilde{Q}_{\Omega}^{\beta}du_{\beta} \quad \ldots \ldots \quad \tilde{g}'_{\Omega\Lambda} = \tilde{Q}_{\Omega}^{\beta}a_{\beta\alpha}\tilde{Q}_{\Lambda}^{\alpha} \; . \tag{61a,b}$$

Here again, the case of $d\tilde{x}'$ written in terms of dx' is included in (61a,b).

-55-

In such a case we would write

$$d\tilde{x}'_\Omega = \tilde{K}^s_\Omega dx'_s \quad \ldots\ldots \quad \tilde{g}'_{\Omega\Lambda} = \tilde{K}^s_\Omega g'_{sr}\tilde{K}^r_\Lambda \ . \tag{61'a,b}$$

According to (59a), equation (61a) is equivalent to (61'a) with

$$\tilde{Q}^\beta_\Omega = Q^\beta_s\tilde{K}^s_\Omega \ ,$$

which, together with (59b), shows that equation (61b) is equivalent to (61'b).

Equations (60'a,b) and (61'a,b) reveal familiar relationships between the (associated) metric tensors even if none of the latter has the full rank. The development of this case has proceeded, in principle, via the (associated) metric tensors for the model surface, which are assumed to have the full rank u.

-56-

4.2 Traditional Identification

In this identification the contravariant components of vectors correspond to the "regular" vectors in adjustment calculus. Although a L.S. adjustment is not yet considered, some of the correspondences below are made in anticipation of the notations used in the following chapter. In identifying second-order mixed tensors and tensor-related quantities with matrices, the contravariant index indicates the row and the covariant index indicates the column. This is compatible with the convention used in adjustment calculus and applied here to $A_\alpha^r \equiv \partial x^r / \partial u^\alpha$, see e.g. equation (2.23) and the statement pertaining to the Jacobian on page 16 of [12]. With regard to other second-order tensors, contravariant or covariant, the first and second indices refer to rows and columns, respectively.

In view of (54a)-(61'b), the correspondences between the tensor and the adjustment notations in the traditional identification are

$$
\left.
\begin{array}{llll}
dx^r & \ldots d\ell \ , & g^{rs} & \ldots C \ , \\[2ex]
du^\alpha & \ldots d\hat{x} \ , & a^{\alpha\beta} & \ldots C_{\hat{x}} \ , \\[2ex]
dx'^r & \ldots d\hat{\ell} \ , & g'^{rs} & \ldots C_{\hat{\ell}} \ ; \\[3ex]
d\bar{x}^i & \ldots d\hat{y} \ , & \bar{g}^{ij} & \ldots C_{\hat{y}} \ , \\[2ex]
d\tilde{x}'^\Lambda & \ldots d\hat{f} \ , & \tilde{g}'^{\Lambda\Omega} & \ldots C_{\hat{f}} \ .
\end{array}
\right\}
\qquad (62)
$$

$$dx_s \quad \ldots \ldots \quad d\ell* \, , \qquad g_{sr} \quad \ldots \ldots \quad P \, ,$$

$$du_\beta \quad \ldots \ldots \quad d\hat{x}* \, , \qquad a_{\beta\alpha} \quad \ldots \ldots \quad P_{\hat{x}} \, ,$$

$$dx'_s \quad \ldots \ldots \quad d\hat{\ell}* \, , \qquad g'_{sr} \quad \ldots \ldots \quad P_{\hat{\ell}} \, ; \qquad \left. \right\} \quad (63)$$

$$d\overline{x}_m \quad \ldots \ldots \quad d\hat{y}* \, , \qquad \overline{g}_{mk} \quad \ldots \ldots \quad P_{\hat{y}} \, ,$$

$$d\tilde{x}'_\Omega \quad \ldots \ldots \quad d\hat{f}* \, , \qquad \tilde{g}'_{\Omega\Lambda} \quad \ldots \ldots \quad P_{\hat{f}} \, .$$

The symbol "*" has been introduced to indicate that the covariant, rather than contravariant, components describe certain adjustment vectors. When dealing with the weight matrices this symbol is no longer necessary (the contravariant or covariant origins are distinguished by the letters C or P, respectively).

The additional correspondences introduced in this section are

$$A^r_\alpha \quad \ldots \ldots \quad A \, ,$$
$$\left. \right\} \quad (64)$$
$$Q^\alpha_r \quad \ldots \ldots \quad Q \, ;$$

$$D^i_r \quad \ldots \ldots \quad D \, , \qquad \tilde{A}^\Lambda_\alpha \quad \ldots \ldots \quad \tilde{A} \, , \qquad \tilde{H}^\Lambda_r \quad \ldots \ldots \quad \tilde{H} \, ; \qquad (65)$$

$$E^s_m \quad \ldots \ldots \quad E \, , \qquad \tilde{Q}^\beta_\Omega \quad \ldots \ldots \quad \tilde{Q} \, , \qquad \tilde{K}^s_\Omega \quad \ldots \ldots \quad \tilde{K} \, . \qquad (66)$$

Starting with the relationships involving the contravariant vector components in all four situations, the formulas of Section 4.1 can now be written as follows:

$$d\hat{y} = Dd\ell \qquad \dots\dots \quad C_{\hat{y}} = DCD^T \,,$$

$$d\hat{x} = Qd\hat{\ell} = Qd\ell \quad \dots\dots \quad C_{\hat{x}} = QC_{\hat{\ell}}Q^T = QCQ^T \,,$$

$$d\hat{\ell} = Ad\hat{x} \qquad \dots\dots \quad C_{\hat{\ell}} = AC_{\hat{x}}A^T \,;$$

$$d\hat{f} = \tilde{A}d\hat{x} \qquad \dots\dots \quad C_{\hat{f}} = \tilde{A}C_{\hat{x}}\tilde{A}^T \,,$$

$$d\hat{f} = \tilde{H}d\hat{\ell} \qquad \dots\dots \quad C_{\hat{f}} = \tilde{H}C_{\hat{\ell}}\tilde{H}^T \,.$$

$$\tag{67}$$

The connection between the last three formulas is provided by

$$\tilde{H}A = \tilde{A} \,.$$

All of the formulas in (67) express the variance-covariance propagation law known from adjustment calculus.

In continuing with the relationships involving the covariant vector components in all four situations, the formulas from Section 4.1 are written as

$$d\hat{y}^* = E^Td\ell^* \qquad \dots\dots \quad P_{\hat{y}} = E^TPE \,,$$

$$d\hat{x}^* = A^Td\hat{\ell}^* = A^Td\ell^* \quad \dots\dots \quad P_{\hat{x}} = A^TP_{\hat{\ell}}A = A^TPA \,,$$

$$d\hat{\ell}^* = Q^Td\hat{x}^* \qquad \dots\dots \quad P_{\hat{\ell}} = Q^TP_{\hat{x}}Q \,;$$

$$d\hat{f}^* = \tilde{Q}^Td\hat{x}^* \qquad \dots\dots \quad P_{\hat{f}} = \tilde{Q}^TP_{\hat{x}}\tilde{Q} \,,$$

$$d\hat{f}^* = \tilde{K}^Td\hat{\ell}^* \qquad \dots\dots \quad P_{\hat{f}} = \tilde{K}^TP_{\hat{\ell}}\tilde{K} \,.$$

$$\tag{68}$$

Here the connection between the last three formulas is provided by

$$Q\tilde{K} = \tilde{Q} \ .$$

The formulas in (68) exhibit exactly the same structure as (67), this time in conjunction with the weight matrices. Accordingly, the feature described by (68) could be called the "weight propagation law".

4.3 New Identification

In the new identification, the "regular" adjustment vectors are described by the covariant components of the corresponding first-order tensors. The indices of second-order tensors and tensor-related quantities follow the convention outlined at the beginning of Section 4.2. The new identification entails the following correspondences:

$$
\left.
\begin{aligned}
dx_s &\ \ldots\ d\ell \ , & g_{sr} &\ \ldots\ C \ , \\[2ex]
du_\beta &\ \ldots\ d\hat{x} \ , & a_{\beta\alpha} &\ \ldots\ C_{\hat{x}} \ , \\[2ex]
dx'_s &\ \ldots\ d\hat{\ell} \ , & g'_{sr} &\ \ldots\ C_{\hat{\ell}} \ ; \\[3ex]
d\overline{x}_m &\ \ldots\ d\hat{y} \ , & \overline{g}_{mk} &\ \ldots\ C_{\hat{y}} \ , \\[2ex]
d\widetilde{x}'_\Omega &\ \ldots\ d\hat{f} \ , & \widetilde{g}'_{\Omega\Lambda} &\ \ldots\ C_{\hat{f}} \ .
\end{aligned}
\right\} \qquad (69)
$$

$$
\left.
\begin{aligned}
dx^r &\ \ldots\ d\ell* \ , & g^{rs} &\ \ldots\ P \ , \\[2ex]
du^\alpha &\ \ldots\ d\hat{x}* \ , & a^{\alpha\beta} &\ \ldots\ P_{\hat{x}} \ , \\[2ex]
dx'^r &\ \ldots\ d\hat{\ell}* \ , & g'^{rs} &\ \ldots\ P_{\hat{\ell}} \ ; \\[3ex]
d\overline{x}^i &\ \ldots\ d\hat{y}* \ , & \overline{g}^{ij} &\ \ldots\ P_{\hat{y}} \ , \\[2ex]
d\widetilde{x}'^\Lambda &\ \ldots\ d\hat{f}* \ , & \widetilde{g}'^{\Lambda\Omega} &\ \ldots\ P_{\hat{f}} \ .
\end{aligned}
\right\} \qquad (70)
$$

The symbol "*" now indicates the use of contravariant components in identifying certain adjustment vectors. As in Section 4.2, this symbol is no longer necessary when dealing with the weight matrices.

It will prove expedient to introduce new correspondences also with regard to A_α^r and other second-order tensors and tensor-related quantities appearing in (64)-(66). The matrix notations A, Q, etc., are again used, but they should be considered completely independent from the matrices which happened to have the same symbol in Section 4.2. The new correspondences are

$$A_\alpha^r \ \ldots\ldots\ Q^T \ , \ \left.\begin{array}{c} \\ \\ \\ \end{array}\right\}$$

$$Q_r^\alpha \ \ldots\ldots\ A^T \ ; \tag{71}$$

$$D_r^i \ \ldots\ldots\ E^T \ , \qquad \tilde{A}_\alpha^\Lambda \ \ldots\ldots\ \tilde{Q}^T \ , \qquad \tilde{H}_r^\Lambda \ \ldots\ldots\ \tilde{K}^T \ , \tag{72}$$

$$E_m^S \ \ldots\ldots\ D^T \ , \qquad \tilde{Q}_\Omega^\beta \ \ldots\ldots\ \tilde{A}^T \ , \qquad \tilde{K}_\Omega^S \ \ldots\ldots\ \tilde{H}^T \ . \tag{73}$$

We note that A_α^r , etc., in the above correspondences could have been written equally well as A_β^S , etc., since in relating these quantities to matrices the role of the indices is merely to indicate the rows and columns.

One can now transcribe the formulas from Section 4.1, starting with the relations involving the covariant vector components in all four situations and continuing with the relations involving the contravariant vector components in all four situations. In the former case the derived formulas would read exactly as (67), including the connecting relation that followed. And in the latter case the formulas would read exactly as (68),

-62-

again including the connecting relation. Thus, the convenient choice
of the matrix notations above has helped to save space, while confirming
the laws of variance-covariance propagation and of "weight propagation"
also in the new identification. This confirmation represents an
important step in establishing the duality between the contravariant
and covariant "spaces" in the L.S. problems as set forth in Chapter 1.

4.4 Rank Considerations

At this stage of the development, it might be instructive to add a few remarks concerning the rank of matrices corresponding to certain second-order tensors. In order to present the results independently of the two kinds of identification, the "regular" adjustment vectors and matrices are represented by the corresponding tensor quantities in brackets (the indices lose much of their role in such cases). First of all, the ranks of g^{rs} and g_{sr} are considered full, i.e., equal to n. This is immediately evident from (15) and (14), respectively. In particular, from (15) it follows that

$$[g^{rs}] = RR^T , \qquad R = [\ [\ell^r] \ [j^r] \ \dots \ [\nu^r] \ \dots \ ] . \qquad (74a,b)$$

Thus, the matrix R of dimensions (n×n) is composed of the contravariant components of the orthonormal vectors ℓ, j, ..., ν, ... and has the full rank n. Accordingly, the (n×n) matrix in (74a) is positive-definite (symmetric). Similar reasoning reveals that

$$[g_{sr}] = SS^T$$

is also positive-definite, with S having the same structure as R except that the covariant vector components are used.

From (33b) we deduce

$$[g'^{rs}] = R'R'^T , \qquad R' = [\ [\ell^r] \ [j^r] \ \dots \ ] , \qquad (75a,b)$$

-64-

where R' of dimensions (n×u) is composed of the contravariant components of the orthonormal vectors ℓ, j, ... , and has the full (column) rank u. Accordingly, the (n×n) matrix in (75a) is positive semi-definite of rank u. In recalling (33a), we reach a similar conclusion also for

$$[g'_{sr}] = S'S'^T ,$$

where S' is formed as R', but in terms of the covariant components.

For the sake of completeness, one can also consider (42b) and deduce

$$[g''^{rs}] = R''R''^T , \qquad R'' = [\ [v^r]\ ... \ ] , \qquad (76a,b)$$

where R'' of dimensions (n×r) has the full (column) rank r. Thus, the (n×n) matrix in (76a) is positive semi-definite of rank r. As before, a similar conclusion in conjunction with (42a) holds also for

$$[g''_{sr}] = S''S''^T .$$

We can accordingly write in analogy to (43b,a):

$$\text{rank}[g^{rs}] = \text{rank}[g'^{rs}] + \text{rank}[g''^{rs}] = u + r = n , \qquad (77)$$

$$\text{rank}[g_{sr}] = \text{rank}[g'_{sr}] + \text{rank}[g''_{sr}] = u + r = n . \qquad (78)$$

Using the same technique, from (20b) we deduce

$$[a^{\alpha\beta}] = MM^T , \qquad M = [\ [\ell^\alpha]\ [j^\alpha]\ ... \ ] , \qquad (79a,b)$$

where M of dimensions (u×u) has the full rank u, making the (u×u) matrix in (79a) positive-definite. Equation (20a) yields

$$[a_{\beta\alpha}] = NN^T ,$$

which has the same dimensions and rank as the matrix in (79a) because N
has the same structure as M except that the covariant vector components
are used. With the newly introduced matrix notations (R, ... , N) we
can write from (27) and (30):

$$[A_\alpha^r] = R'N^T ,$$ (80)

$$[Q_r^\alpha] = MS'^T .$$ (81)

Since N, M are square matrices of full rank, the rank of the (n×u) matrix
in (80) equals that of R' (namely u) and the rank of the (u×n) matrix in
(81) equals that of S' (also u). We can thus summarize the rank results
for the last four entities as

$$\text{rank}[a^{\alpha\beta}] = \text{rank}[a_{\beta\alpha}] = \text{rank}[A_\alpha^r] = \text{rank}[Q_r^\alpha] = u .$$ (82)

As a matter of interest, we note from (36) and (37) that

$$\text{trace}[A_\alpha^r Q_t^\alpha] = \text{trace}[Q_r^\alpha A_\beta^r] = u ,$$

due to the property that trace(AQ)=trace(QA), where QA is the unit matrix
of dimensions (u×u) as is seen from (36). In the same vein, we have

$$\text{trace}[S_L^r B_t^L] = \text{trace}[B_r^L S_M^r] = r .$$

5. LEAST-SQUARES ADJUSTMENT IN TRADITIONAL IDENTIFICATION

5.1 Parametric Method

The underlying mathematical model for the parametric method can be expressed as

$$\ell = F(x) , \tag{83}$$

where, in the adjustment notations, ℓ is the vector of n observables, x is the vector of u parameters and F represents the relationship between the two sets. The model (83) is occasionally thought of as relating some "true" quantities ℓ and x, in which case the superscript "t" can be used. Although the philosophy of utilizing the "true" values may be questioned, the usefulness of the model (83) is safeguarded by virtue of the consideration that the adjusted values must satisfy the same relationships. In this context the symbol "^" replaces the superscript "t". And when the same model relates some initial values, the superscript "°" is used. Clearly, the observables ℓ° (constants) are consistent with the initial parameters x°. This could not be said of the actual observations denoted here as $\tilde{\ell}$, which in general (for n>u) are not consistent with any set of parameters.

Considering the model (83) in terms of the adjusted quantities, we simplify it upon using only the linear terms of the Taylor expansion; the usual assumptions regarding good initial values of the parameters, etc., can be found in standard textbooks on the subject and need not be repeated here. The linearized adjustment model reads

$$\hat{\ell} = A d\hat{x} + \ell^\circ , \tag{84}$$

-67-

where

$$A \equiv (\partial F/\partial x)_{x^\circ} \; , \quad d\hat{x} \equiv \hat{x} - x^\circ \; , \quad \ell^\circ \equiv F(x^\circ) \; . \quad\quad (85a,b,c)$$

The adjusted values $\hat{\ell}$ result from the observed values $\tilde{\ell}$ minus the error estimates denoted as $\hat{\epsilon}$; the $-\hat{\epsilon}$ are often called (plus) corrections, or residuals. Equation (84) can be transcribed as

$$d\ell = d\hat{\ell} + \hat{\epsilon} \; , \quad\quad\quad\quad\quad\quad (86)$$

where

$$d\ell = \tilde{\ell} - \ell^\circ \; , \quad d\hat{\ell} \equiv \hat{\ell} - \ell^\circ = Ad\hat{x} \; , \quad \hat{\epsilon} = \tilde{\ell} - \hat{\ell} \; . \quad\quad (87a,b,c)$$

A more usual transcription of (84) in adjustment calculus would read

$$-\hat{\epsilon} = Ad\hat{x} + (\ell^\circ - \tilde{\ell}) \; ,$$

where $-\hat{\epsilon}$ could also be denoted as V, $d\hat{x}$ as X, and $\ell^\circ - \tilde{\ell}$ as $L^\circ - L^b \equiv L$. However, the equivalent formulation (86)-(87c) is more readily adaptable to the geometrical interpretation.

By using similar notations in a more explicit adjustment formulation which we are about to explain, we are led to a geometrical analogue of the above equations in a simple and natural manner. The main feature of these notations is that the vectors in (83) are written in terms of their individual components which, in anticipation of the traditional identification, are designated by superscripts. We thus write equation (83) as

$$\{\ell^r\} = \{F^r(x^\alpha)\} \; , \quad\quad\quad\quad\quad\quad (88)$$

-68-

where $F^1(x^1,x^2,...,x^u)$, $F^2(x^1,x^2,...,x^u)$, ... , $F^n(x^1,x^2,...,x^u)$ can be regarded as n coordinates $F^r \equiv \ell^r$ expressed in terms of u coordinates x^α. But this corresponds perfectly to Gauss' form of a surface as described prior to equation (25), expressing each space coordinate (here ℓ^r) as some function of the u surface coordinates (here x^α). We can now consider the specific values of surface coordinates, $x^{\circ \alpha}$, as describing a point in the u-dimensional surface, and the corresponding values of space coordinates, $\ell^{\circ r}$, as describing the same point in the underlying n-dimensional space. This point is denoted as P. The transcription of (84)-(87c) with the present notations would be exceedingly simple, in that $d\hat{x}$, $\hat{x}$ and x° would be attributed the superscript "α" and all the other vector quantities would be attributed the superscript "r". From (85a) we observe that the element (r,α) of the matrix A is $\partial F^r/\partial x^\alpha \equiv \partial \ell^r/\partial x^\alpha$ evaluated at the point P. Thus, viewed in the context of coordinates, the partial derivatives in (85a) correspond to the components of the tensor quantity within the parentheses of (25) written in (27) as A_α^r.

The analogy between the setup and solution of the parametric method on one hand, and the geometrical configuration featuring the model surface on the other hand can now be completed. The discussion in the preceding paragraph leads directly to the following correspondences between the adjustment relations (86), (87b), and the tensor formulation of Chapter 3 (equations 38b, 28a):

$$d\ell = d\hat{\ell} + \hat{\varepsilon} \quad ..... \quad dx^r = dx'^r + dx''^r \, , \tag{89a}$$

$$d\hat{\ell} = A d\hat{x} \quad ..... \quad dx'^r = A_\alpha^r du^\alpha \, . \tag{89b}$$

-69-

Since tensors in general are functions of position, the tensor forms
on the right-hand sides of (89a,b) must belong to a specific point.
Clearly, this point is P defined earlier. To see this for (89a) we only
have to notice (87a,b) containing $\ell°$ and implying that the vectors
represented by the components dx^r and dx'^r can indeed be associated with
P. The vector represented by the components dx''^r completes the vector
equation in (89a) and, therefore, can likewise be associated with P. In
three dimensions, such a situation is depicted in Fig. 1. With regard
to (89b), du^α is associated with P through the values $x°$ in (85b) and A has
already been shown to be linked to P.

We notice that the correspondence shown in (89b) could have been arrived
at even without the benefit of the natural correspondence between A and the
partial derivatives $\partial x^r/\partial u^\alpha \equiv A_\alpha^r$. In particular, one could have discarded any
reference to these partial derivatives and defined A_α^r as $\ell^r \ell_\alpha + j^r j_\alpha + \ldots$,
where all of the vectors ℓ, j, ... are considered at P. That this definition
has both the necessary and sufficient properties follows from (22a) and (23a),
based on dx' and du being considered the same physical vector. But such an
approach is merely academic since we know from (27) that the partial deriva-
tives $\partial x^r/\partial u^\alpha$ and A_α^r as just defined are one and the same quantity. However,
a similar vectorized approach will be useful in the new identification,
where the tensor equivalent of A will not be expressed by means of any
partial derivatives. Since the design matrix A in adjustment calculus is
indeed composed of partial derivatives, one is compelled to appreciate the
traditional identification making the geometrical interpretation of the L.S.
formulation exceedingly straightforward.

-70-

The solution of the present L.S. problem is derived from the familiar minimum condition for

$$V^T PV \equiv \hat{\epsilon}^T P \hat{\epsilon} \ . \tag{90}$$

Since the vector $\hat{\epsilon}$ corresponds to the components dx''^r (see equation 89a), from the first equality in (39) we confirm that P, the weight matrix of observations, corresponds to g_{sr}. All the correspondences between the tensor notations in the traditional identification and the adjustment notations can thus be adopted from the first three lines of (62) and from (64). The solution and all the derived relations in the parametric method can then be obtained from the right-hand side of the geometrical diagram in Fig. 2. To facilitate this task, Fig. 3 has been constructed containing the same diagram except that the adjustment notations have replaced their tensor counterparts.

From the right-hand side of the above diagram we can extract most of the standard results for the parametric method. Clearly, all the quantities of interest should be expressed in terms of the known quantities identified by heavy lines. Following the rules explained in Chapter 3, we can write

$$d\hat{x} = Qd\hat{\ell} = Qd\ell \equiv C_{\hat{x}} A^T P d\ell \ , \tag{91}$$

where

$$C_{\hat{x}} = P_{\hat{x}}^{-1} \ , \qquad P_{\hat{x}} = A^T PA \ . \tag{92a,b}$$

The matrix $P_{\hat{x}}$ has been listed in two forms in (68) but only the second form, (92b), is useful in practice. Since Q can only be obtained from $C_{\hat{x}}$, it is

not suitable for actual computations. But if it were expressed explicitly it could be used for a straightforward verification of $C_{\hat{x}}$ in (67). With regard to the adjusted observations, as in (67) we have

$$(93a,b)$$

$$d\hat{\ell} = A d\hat{x} \ (\equiv AQd\hat{\ell} \equiv AQd\ell), \qquad C_{\hat{\ell}} = AC_{\hat{x}}A^T \ (\equiv AQC_{\hat{\ell}}Q^TA^T \equiv AQCQ^TA^T);$$

the equivalent relations in parentheses could be further extended. The (singular) matrix $P_{\hat{\ell}}$, not sought in the standard adjustment, can be expressed as

$$P_{\hat{\ell}} = Q^T P_{\hat{x}} Q \equiv PC_{\hat{\ell}}P \ , \qquad\qquad (93c)$$

whose first form has been listed in (68). Its second form also follows from the diagram. The functions of $d\hat{x}$ or $d\hat{\ell}$, as well as their C-matrices, *can be computed as in the* last two lines of (67). The P-matrices of these functions, although listed in (68), cannot be computed under normal circumstances because $\tilde{Q}$ and $\tilde{K}$ are unknown. They have been listed in Chapter 4 mainly for the purpose of establishing the weight propagation law in various configurations.

The error estimates (or minus the residuals), $\hat{\varepsilon}$, as well as their C- and P-matrices follow from the diagram as

$$\hat{\varepsilon} = d\ell - d\hat{\ell} \ , \qquad C_{\hat{\varepsilon}} = C - C_{\hat{\ell}} \ , \qquad P_{\hat{\varepsilon}} = P - P_{\hat{\ell}} \ . \qquad (94a,b,c)$$

Corresponding to $ds^{"2}$, from (52a) we have

$$\hat{\varepsilon}^T P \hat{\varepsilon} = d\ell^T P d\ell - d\hat{x}^T A^T P d\ell \equiv d\ell^T P d\ell + d\hat{x}^T u \ , \qquad (95a)$$

where

-72-

$$u = -A^T P d\ell \equiv A^T P(\ell° - \tilde{\ell}) \ . \tag{95b}$$

In terms of the notations introduced following (87c), the last term above could be written as $A^T PL$.

5.2 Condition Method

Let us consider the mathematical model $G(\ell)=0$, where G is a set of known functions and ℓ could represent an unspecified number of other functions characteristic of a given problem (for example, the latter could include all possible angles in a triangulation network, etc.). However, in the condition method the set ℓ is limited to the observables as introduced in the parametric method (with regard to the example just mentioned, only the angles slated for actual measurements would now be included). We can thus write

$$G(\ell) = 0 , \qquad \ell = F(x) . \qquad (96a,b)$$

But no matter what kind of parameters x are adopted to properly describe the observables ℓ in (96b), the equality (96a) holds true. Similar to Section 5.1, symbols "t", "^" or "°" can be used in conjunction with ℓ and x. Since equations (96a,b) allow us to view G as (composite) functions of x, we have

$$(\partial G/\partial x)_{x^\circ} = 0 , \qquad (97)$$

where the point of evaluation has been chosen to be x°.

In considering a well-defined problem with n observables and u parameters as in Section 5.1, the number of independent relations (96a) is

$$r = n - u . \qquad (98)$$

In particular, in such a problem there must exist u observables, called "necessary", which can uniquely determine the u parameters. Therefore,

-74-

the remaining r observables must be expressible in terms of the necessary observables. Any one of these r observables minus its counterpart expressed in terms of the necessary observables can be regarded as a new function, equal to zero. The totality of these new functions represents r independent relationships (each involves a different set of r+1 observables in general). We have thus arrived at an independent set of r functions involving all n observables, which can be symbolized by (96a). This heuristic reasoning could be made simpler if only linear functions were considered in (96a,b). It would then become immediately apparent that (96a) can represent no more than r independent functions, etc. However, there is no need to expand this discussion whose sole purpose has been to illustrate the well-known fact (98).

As a consequence of (98), the matrix in (97) has the dimensions $(r \times u)$. In terms of (96a,b) one can express (97) as

$$BA = 0 , \qquad B = (\partial G/\partial \ell)_{\ell^\circ} , \qquad A = (\partial \ell/\partial x)_{x^\circ} , \qquad (99a,b,c)$$

where the matrix B has the dimensions $(r \times n)$ and A, already encountered in (85a), has the dimensions $(n \times u)$. In line with Section 5.1, the partial derivatives (99b,c) are associated with the point P. Since in practice (96b) is usually not utilized in conjunction with the condition method alone, ℓ° is not known and B in (99b) is computed with the actual observations $\tilde{\ell}$ replacing ℓ°. If, in such cases, A could also be computed (and expressed as in 99c), equation (99a) would be confirmed sufficiently well for most purposes, but would not be completely rigorous. This stems from the fact that although (96b) with the values "°" is a consistent system of equations, it is not so

-75-

with $\tilde{\ell}$ replacing ℓ° (i.e., $\tilde{\ell} \neq F(x^\circ)$, see also the first paragraph in Section 5.1). Under these circumstances the matrix product BA is not exactly the same matrix as the one appearing on the left-hand side of (97).

The linearization process via the Taylor expansion at P can be presented as follows:

$$G(\hat{\ell}) = G(\ell^\circ) + (\partial G/\partial \ell)_{\ell^\circ} (\hat{\ell} - \ell^\circ) = 0 , \qquad (100)$$

where (96a) has been used for $G(\hat{\ell})$. In using it also for $G(\ell^\circ)$, and recalling (99b) and (87b), equation (100) becomes

$$B d\hat{\ell} = 0 . \qquad (101)$$

If (86) is taken into account, (101) can in turn be written as

$$B d\ell = B \hat{\epsilon} .$$

If, in the expansion (100), $\hat{\ell}$ is replaced by $\tilde{\ell}$, the result is a vector w (often called "misclosures") rather than zero. Using the notation (87a), in this case we obtain

$$B d\ell = w \equiv G(\tilde{\ell}) .$$

The relation following (101) can thus be completed to read

$$w = B \hat{\epsilon} = B d\ell . \qquad (102)$$

The first equality above resembles the usual setup for the condition method, namely

$$B(-\hat{\epsilon}) + w = 0 . \qquad (103)$$

-76-

If B is evaluated with $\ell°$ we call (103) or, equivalently, (102) the "theoretical" formulation, and if B is evaluated with $\tilde{\ell}$, we call (103) the "practical" formulation. The differences in results between these two formulations are likely to be insignificant. It should be pointed out that the practical formulation entails no more approximations than the theoretical formulation because it is again a consequence of an expansion like (100), except that $\ell°$ is replaced by $\tilde{\ell}$.

We notice that if the values $\ell°$ -- and thus also $d\ell$ and the "theoretical" B -- were available, the misclosures w would be computed as $Bd\ell$ according to (102) and not as $G(\tilde{\ell})$, except for verification purposes. The computational drawback of this theoretical formulation lies in the necessity of introducing parameters in the model, with which (96b) is evaluated at P in analogy to the parametric method. If the process of expressing the observables in terms of parameters proved too cumbersome (especially if u is much larger than r), we could compute the misclosures as $G(\tilde{\ell})$ instead, etc., and disregard the parameters altogether. We would thus be adopting the practical formulation. However, the geometrical approach of Chapter 3 coincides with the theoretical formulation, in that it treats the tensor quantities in both surfaces as belonging to the same point P ($B_r^L A_\alpha^r = 0$ holds exactly, etc.). Accordingly, we shall keep pursuing the theoretical formulation and take advantage of the wealth of information encompassed by the diagram of Fig. 2, including various interrelations between the parametric and condition methods.

At this point one may be compelled to ask: How can we utilize the formulas derived through the theoretical formulation if we wish to proceed, in actual computations, without having to evaluate $\ell°$ and $d\ell$? Clearly, in

-77-

the theoretical formulation $d\ell$ is considered as the basic known quantity rather than w, which has this property in the practical formulation. Fortunately, no essential relations in the theoretical formulation utilize $d\ell$ alone, but only in the combination $Bd\ell$ which is w, the same in both formulations. Other simple correspondences exist. For example, using the theoretical formulation we can compute $d\hat{\ell}$ from $d\ell$ and $\hat{\varepsilon}$, as well as its C-matrix. For the adjusted observations we then have

$$\hat{\ell} = \ell^\circ + d\hat{\ell} = \tilde{\ell} - \hat{\varepsilon} . \tag{104}$$

Although the first equality is of no use in the practical formulation, the second equality serves our purpose just as well. Furthermore, as is the case with $\tilde{\ell}$ and $d\ell$, $\hat{\ell}$ and $d\hat{\ell}$ differ only by a constant vector ℓ° and, therefore, the C-matrix for $d\hat{\ell}$ is also the C-matrix for $\hat{\ell}$. In the same vein, all the final formulas based on the theoretical formulation can find their use in the practical formulation as well.

In view of the upcoming development we note that any "new" formula (96a), obtained from the original one through the pre-multiplication by an arbitrary nonsingular matrix D of dimensions $(r \times r)$, would represent merely another set of r independent relations containing the same amount of information as its original counterpart. The formulas such as (99a), (101), (102) and (103) would hold true also in this new situation, except that the new B and new w would be equal to their original counterparts pre-multiplied by the same matrix D. Since it does not matter which set (96a) of r independent relations constitutes the starting point for the condition method, any matrix B can be used in the theoretical formulation (the

-78-

appropriate w then follows from 102), provided it is obtained from some "original" B through the pre-multiplication by D as defined above. We need not distinguish between such "original" and derived matrices B. If one is more suitable than another in the derivations, we can simply adopt it and denote it by the same letter "B".

In relating the adjustment setup for the condition method (theoretical formulation) to the geometrical configuration of Chapter 3, we focus our attention on equation (101). Since in Section 5.1 $d\hat{\ell}$ has already been identified with the components dx'^r of the vector dx' lying in the model surface, it follows that the r "rows" of the desired second-order tensor to be identified with B must be formed by the space covariant components of vectors q orthogonal to the model surface (hence lying in the second surface). These vectors q are mutually independent due to the full (row) rank of B, and must therefore be expressible as independent combinations involving all of the r orthonormal vectors ν, ... In the matrix form, the desired second-order tensor can be written as KT, where the nonsingular matrix K of dimensions (r×r) contains the coefficients of the "independent combinations" above, and the matrix T of dimensions (r×n) has as its rows the space covariant components of ν, ... The following matrix D is admissible to modify the "original" matrix B:

$$D = [\ [\nu^{\llcorner}]\ \ldots\ ]K^{-1}\ ,$$

where $[\nu^{\llcorner}]$, ... are the column-vectors of r elements each, representing the surface contravariant components of ν, ... Therefore, the desired second-order tensor written in the matrix form is the product of two matrices,

where the first (of dimensions r×r) has the components v^L, ... as its columns, and the second (of dimensions r×n) has the components v_r , ... as its rows.

As a result of the above development, the (L,r) element of the matrix representing the desired tensor is $v^L v_r +...$ This tensor is accordingly B^L_r as defined in (41). In analogy to (89a,b), we can now write two basic correspondences for the condition method:

$$d\ell = d\hat{\ell} + \hat{\epsilon} \quad ..... \quad dx^r = dx'^r + dx''^r , \qquad (105a)$$

$$Bd\hat{\ell} = 0 \quad ..... \quad B^L_r dx'^r = 0 , \qquad (105b)$$

where (105a) is the same as (89a). All the tensor quantities are again associated with the point P.

In employing the reasoning and identifications parallel to those in (90) and the text that followed, we can write the correspondences with regard to the second surface and the condition method as

$$\left. \begin{array}{llll} dx^r & ..... & d\ell , & g^{rs} & ..... & C , \\ du^L & ..... & w , & a^{LM} & ..... & C_w , \\ dx''^r & ..... & \hat{\epsilon} , & g''^{rs} & ..... & C_{\hat{\epsilon}} ; \end{array} \right\} \qquad (106)$$

$$\left. \begin{array}{llll} dx_s & ..... & d\ell^* , & g_{sr} & ..... & P , \\ du_M & ..... & w^* , & a_{ML} & ..... & P_w , \\ dx''_s & ..... & \hat{\epsilon}^* , & g''_{sr} & ..... & P_{\hat{\epsilon}} ; \end{array} \right\} \qquad (107)$$

-80-

$$B_r^L \quad \ldots \ldots \quad B \, ,$$

$$S_L^r \quad \ldots \ldots \quad S \, . \qquad \qquad \Big\} \qquad \qquad (108)$$

These correspondences are reflected on the left-hand side of the diagram in Fig. 3.

The diagram of Fig. 3 leads us to appreciate the complete analogy between the parametric and condition methods, whose final formulas differ due to the dissimilar matrices known a priori (A versus B) and not to some differences in their structure. The structural analogies between the two methods are rooted in the analogies between their geometrical counterparts in Fig. 2, the model surface and the second surface. Thus, the misclosures w in the condition method correspond to the parameters $d\hat{x}$ in the parametric method. This is confirmed by comparing equations (91) and (102). And as in the parametric method, no Lagrange multipliers are needed in the present geometrical representation of the condition method, either. It will become apparent that the Lagrange multipliers used in the standard resolution of the condition method correspond to -w* in Fig. 3, offering little insight or advantage in the present context.

The left-hand side of the diagram in Fig. 3 enables us to extract most of the standard results for the condition method. For w regarded as parameters we have

$$w = Bd\ell \, , \qquad C_w = BCB^T \, , \qquad \qquad (109a,b)$$

where (109a) already appeared in (102), and where C_w is sometimes denoted

as M in the standard adjustment approach. For the sake of interest we also present

$$w^* = P_w w , \qquad P_w = C_w^{-1} , \tag{110a,b}$$

where the formulation of P_w through S has been avoided. Except for the sign, w^* above agrees with the standard formula giving the Lagrange multipliers as $K = -(BCB^T)^{-1}w$. The functions of w or $\hat{\epsilon}$, as well as their C-matrices, could be expressed using the appropriate "transliteration" of the last two lines in (67). However, such quantities are not likely to be needed in practice.

The error estimates $\hat{\epsilon}$ and their C- and P-matrices are deduced from the diagram as

$$\hat{\epsilon} = Sw \equiv CB^T P_w w \ (\equiv SB\hat{\epsilon} \equiv SBd\ell \equiv CP_{\hat{\epsilon}}d\ell), \tag{111a}$$

$$C_{\hat{\epsilon}} = SC_w S^T \equiv CB^T P_w BC \ (\equiv SBC_{\hat{\epsilon}}B^T S^T \equiv SBCB^T S^T \equiv CP_{\hat{\epsilon}}C), \tag{111b}$$

$$P_{\hat{\epsilon}} = B^T P_w B . \tag{111c}$$

In expressing the L.S. quadratic form as $ds^{"2}$ in (53a), we have

$$\hat{\epsilon}^T P \hat{\epsilon} = w^T P_w w \equiv w^{*T} w , \tag{112}$$

which, in the standard L.S. approach, can be written as $-K^T w$.

A number of "cross-products" can be read from the diagram. In the form matrix-vector we have, for example,

$$C_{\hat{\ell}}\hat{\epsilon}^* = 0 , \qquad P_{\hat{\ell}}\hat{\epsilon} = 0 \ldots d\hat{\ell}^T P_{\hat{\ell}}\hat{\epsilon} \equiv d\hat{\ell}^T P \hat{\epsilon} \equiv d\hat{\ell}^{*T}\hat{\epsilon} = 0 ,$$
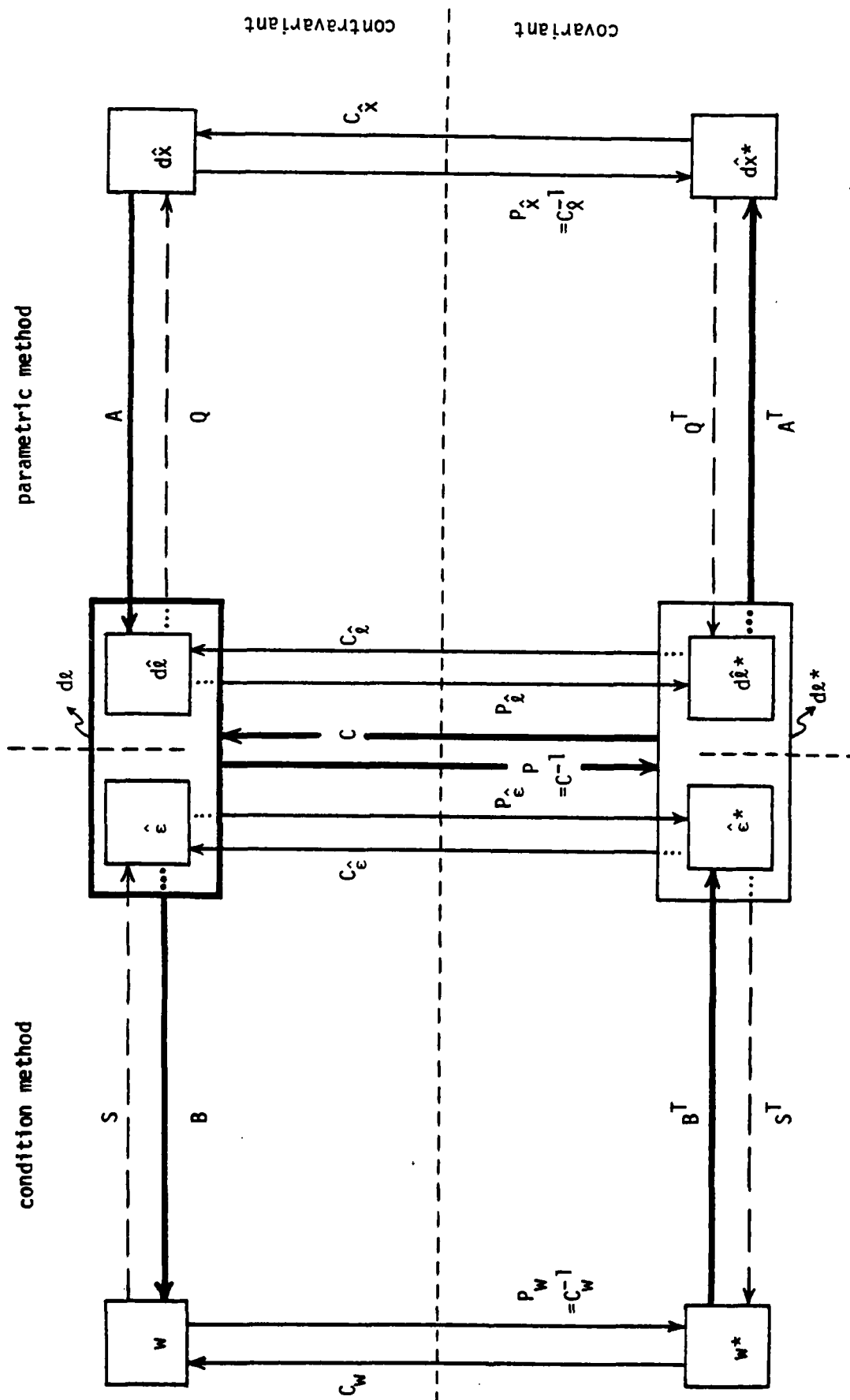
Figure 3

Least-squares commutative diagram (traditional identification)

and several other products including their symmetric analogues (with respect to the vertical dashed line). One of the more important relations of this kind, used as a means of verification in the parametric method, is

$$A^T \hat{\epsilon}^* \equiv A^T P \hat{\epsilon} = 0 \ . \tag{113}$$

In the theoretical formulation one would also have $Bd\hat{\ell}=0$; however, in the practical formulation $d\hat{\ell}$ would not be computed. With regard to the matrix-matrix form, one can readily transcribe (48a-f), for example, of which the first three relations now read

$$BA = 0 \ , \quad QS = 0 \ , \quad QC_{\hat{\ell}}B^T \equiv QCB^T = 0 \ , \tag{114a,b,c}$$

where (114a) has already appeared as (99a). We could similarly list the zero products along the vertical lines, such as between $P_{\hat{\epsilon}}$ and $C_{\hat{\ell}}$, etc. A useful verification formula can be obtained using "cross-additions" such as (50), whose counterpart in matrix notations is

$$C_{\hat{\ell}}P_{\hat{\ell}} + C_{\hat{\epsilon}}P_{\hat{\epsilon}} = AQ + SB = CP = I \ . \tag{115a}$$

In expressing the matrices Q and S from the diagram, (115a) yields

$$AC_{\hat{x}}A^T P + CB^T P_w B = I \ . \tag{115b}$$

In terms of adjustment notations, the statement below (38a,b) can be interpreted in the sense that AQ acts as an orthogonal projection operator from the space onto the model surface (see equation 93a, second expression inside the parentheses), and that SB acts as an orthogonal projection operator from the space onto the second surface (see equation 111a, second

-84-

expression inside the parentheses). From equations (114a,b) above we readily confirm that these two operators are orthogonal to each other. And equation (115a) confirms that their sum yields the identity operator. Clearly, these as well as other L.S. formulas and properties implied in this chapter (whether stated explicitly or not) are a natural rewrite of their geometrical predecessors presented in Chapter 3 and a part of Chapter 4.

6. LEAST-SQUARES ADJUSTMENT IN NEW IDENTIFICATION

6.1 Parametric Method

In the new identification, the basic correspondences between the adjustment and tensor notations can be developed in analogy to (89a,b) with two exceptions: the vector components are now covariant (one can imagine the contravariant index r changed to the covariant index s), and the tensor in (89b) is to be replaced by a different second-order tensor. Since $\hat{\varepsilon}$ in this identification corresponds to dx''_s, from (39) we confirm that the weight matrix P corresponds to g^{rs}. The task before us is thus limited to finding the second-order tensor relating dx'_s and du_β, where dx' and du represent the same physical vector as before. This relationship should correspond to the adjustment formula $d\hat{\ell}=Ad\hat{x}$. It is already clear that the desired tensor has β for its contravariant index and s for its covariant index. Since in the matrix form β designates the rows and s designates the columns according to the previous definition (see Section 4.2 and the mention of the reference [12]), this tensor could be associated with the adjustment matrix A in a direct manner only upon changing the rules for matrix multiplication when dealing with the new identification. A much more palatable procedure is to leave the multiplication rules intact and associate the desired tensor with A^T instead. Under such circumstances, the matrix product $Ad\hat{x}$ will be expressed by the transpose of the desired tensor in the matrix form, post-multiplied by $[du_\beta]$, the column-vector composed of the elements du_β.

-86-

In order to relate dx'_s and du_β, equations (22b) and (23b) are consulted yielding $a = dx'_s \ell^s = du_\beta \ell^\beta$, etc., from which it follows that

$$dx'_s = (\ell^\beta \ell_s + j^\beta j_s + \dots)du_\beta \equiv Q_s^\beta du_\beta .\qquad (116)$$

This result has already been presented as (31a). The desired second-order tensor corresponding to A^T is thus seen to be Q_s^β. Similarly, *the adjustment formula* (91) corresponds now to (44a), from which one can deduce that the tensor A_β^s is associated with Q^T. One can summarize the correspondences pertaining to the model surface by rewriting the first three lines of (69), the first three lines of (70), and (71):

$$\left.\begin{array}{ll} dx_s \ \dots\dots \ d\ell , & g_{sr} \ \dots\dots \ C , \\[2ex] du_\beta \ \dots\dots \ d\hat{x} , & a_{\beta\alpha} \ \dots\dots \ C_{\hat{x}} , \\[2ex] dx'_s \ \dots\dots \ d\hat{\ell} , & g'_{sr} \ \dots\dots \ C_{\hat{\ell}} ; \end{array}\right\} \qquad (117)$$

$$\left.\begin{array}{ll} dx^r \ \dots\dots \ d\ell* , & g^{rs} \ \dots\dots \ P , \\[2ex] du^\alpha \ \dots\dots \ d\hat{x}* , & a^{\alpha\beta} \ \dots\dots \ P_{\hat{x}} , \\[2ex] dx'^r \ \dots\dots \ d\hat{\ell}* , & g'^{rs} \ \dots\dots \ P_{\hat{\ell}} ; \end{array}\right\} \qquad (118)$$

$$\left.\begin{array}{l} A_\alpha^r \ \dots\dots \ Q^T , \\[2ex] Q_r^\alpha \ \dots\dots \ A^T . \end{array}\right\} \qquad (119)$$

With the above correspondences the right-hand side of the diagram in Fig. 2 can be used to construct the right-hand side of the new diagram presented in Fig. 4. In this process, (119) is to be understood in the sense that if A_γ^t is involved in a contraction with a tensor whose (contravariant) index is γ, it corresponds to Q^T; and if it is contracted with a tensor whose (covariant) index is t, it corresponds to Q. The tensor Q_t^γ is treated similarly with self-evident changes. Finally, the arrows designating A and A^T are drawn in heavy lines indicating the quantities known a priori, while the arrows designating Q and Q^T (resulting from products of other matrices) are drawn in dashed lines. If the right-hand sides of Figs. 3 and 4 are compared, it is apparent that not only are all the relationships identical, but these two parts are essentially mirror images of each other along the dashed horizontal line. As well, the error estimates and their C- and P-matrices in the new identification are identical to (94a,b,c), and the L.S. quadratic form expressed with the aid of (52b) and the selected correspondences from (117)-(119) is identical to (95a,b).

6.2 Condition Method

An examination of this method in the new identification reveals that two changes are needed in the correspondences of (105a,b): the vector components are now covariant (the upper index r changes to the lower index s), and the tensor in (105b) is to be replaced by a different second-order tensor whose contravariant and covariant indices are s and M, respectively. Using similar reasoning to that of the first paragraph in Section 6.1, we conclude that this desired tensor will correspond to the adjustment matrix B^T. In order to find it we can proceed in a loose analogy to the part of Section 5.2 dealing with the quasi-arbitrary matrix D. In view of the adjustment formula $Bd\hat{\ell}=0$, the desired tensor in the transposed matrix form post-multiplied by $[dx_s']$ equals the zero vector. Accordingly, the rows of this transposed matrix consist of the space contravariant components of vectors q orthogonal to the model surface (hence lying in the second sur-face). Similar to the procedure used earlier, the vectors q can be expres-sed by independent combinations involving all of ν, ... , and thus the above transposed matrix can be written as K'T', where K' is a nonsingular matrix of dimensions (r×r) and T' is a matrix of dimensions (r×n) having as its rows the space contravariant components of ν, ... In using $[[\nu_M]...]{K'}^{-1}$ as the D-matrix to modify the "original" B, the modified form of the desired transposed matrix is the product of two matrices, where the first (of dimensions r×r) has the components ν_M , ... as its columns and the second (of dimensions r×n) has the components ν^s, ... as its rows.

The above development implies that the (M,s) element of the desired

transposed matrix is $\nu_M \nu^S +\ldots$ Therefore, the desired tensor itself is S_M^S according to the definition (40). Next we consider the adjustment formula (111a) and seek a tensor corresponding to the matrix S. Since the adjustment vectors $\hat{\epsilon}$ and w correspond now to the components dx_S'' and du_M (representing the same physical vector), respectively, we are seeking a second-order tensor which would relate the latter. In analogy to (116), we find it to be $\nu^M \nu_S +\ldots$, or B_S^M, according to the definition (41). The same considerations as before imply that this tensor corresponds to the adjustment matrix S^T. In the new identification, the correspondences paralleling (106)-(108) are listed as

$$
\left.
\begin{array}{llll}
dx_s & \ldots\ldots\ d\ell \ , & g_{sr} & \ldots\ldots\ C\ , \\[2ex]
du_M & \ldots\ldots\ w\ , & a_{ML} & \ldots\ldots\ C_w\ , \\[2ex]
dx_s'' & \ldots\ldots\ \hat{\epsilon}\ , & g_{sr}'' & \ldots\ldots\ C_{\hat{\epsilon}}\ ;
\end{array}
\right\} \tag{120}
$$

$$
\left.
\begin{array}{llll}
dx^r & \ldots\ldots\ d\ell^\star\ , & g^{rs} & \ldots\ldots\ P\ , \\[2ex]
du^L & \ldots\ldots\ w^\star\ , & a^{LM} & \ldots\ldots\ P_w\ , \\[2ex]
dx''^r & \ldots\ldots\ \hat{\epsilon}^\star\ , & g''^{rs} & \ldots\ldots\ P_{\hat{\epsilon}}\ ;
\end{array}
\right\} \tag{121}
$$

$$
\left.
\begin{array}{ll}
B_r^L & \ldots\ldots\ S^T\ , \\[2ex]
S_L^r & \ldots\ldots\ B^T\ .
\end{array}
\right\} \tag{122}
$$
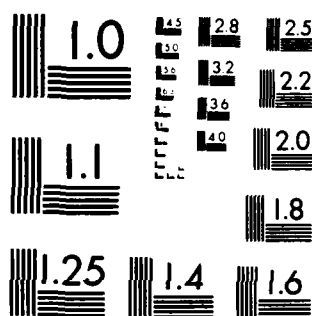
MICROCOPY RESOLUTION TEST CHART
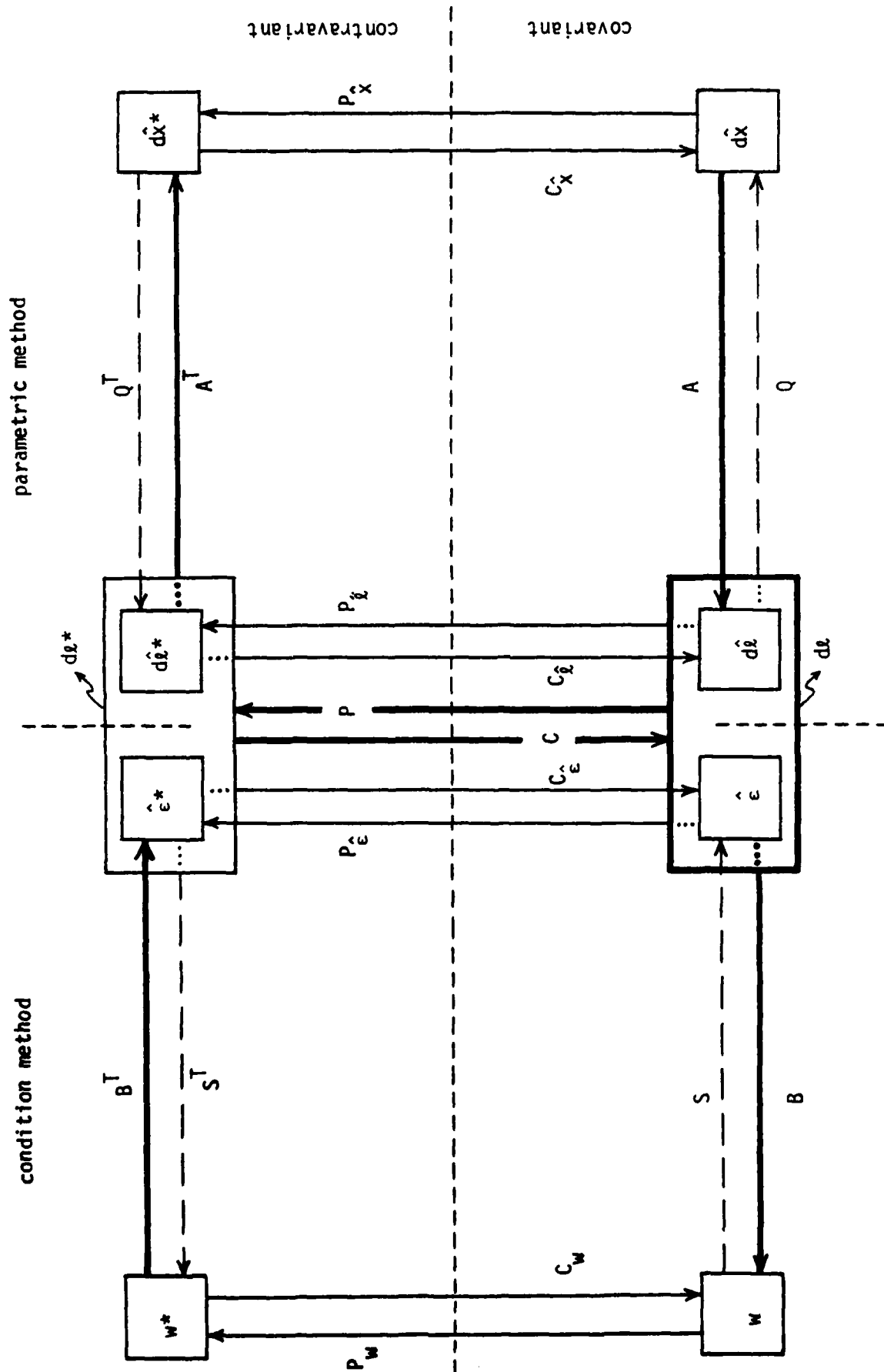NATIONAL BUREAU OF STANDARDS-1963 A

Figure 4

Least-squares commutative diagram (new identification)

With these correspondences the left-hand side of the diagram in Fig. 2 can be used to complete the diagram of Fig. 4. The interpretation of (122) is similar to the interpretation of (119) described earlier. The arrows designating B and B^T in Fig. 4 are drawn in heavy lines while the arrows designating S and S^T are drawn in dashed lines, the reasons being again similar to those given in Section 6.1. The comment regarding the mirror images can likewise be repeated here, indicating that Figs. 3 and 4 in their entirety have this property. The error estimates and their C- and P-matrices are necessarily the same as (111a,b,c), being just three of all possible relationships identical between Figs. 3 and 4. Likewise, the L.S. quadratic form expressed with the aid of (53b) and the selected correspondences from (120), (121) is identical to (112).

7. CONCLUSIONS

Perhaps the most important outcome of this development has been the
geometrical commutative diagram of Fig. 2 built through the extensive
application of tensor algebra. By adapting the geometrical configuration
in a natural way (via the traditional identification) to the adjustment
setup, we have been able to construct the commutative diagram of Fig. 3
expressing most of the standard L.S. formulas as well as several new
relationships. Using another adaptation (via the new identification), we
have constructed the commutative diagram of Fig. 4 expressing all the
pertinent relationships as identical to those given by Fig. 3. In fact,
the two diagrams are essentially mirror images of each other. This
demonstrates the complete equivalence between the two identifications
which, in turn, underlines the possibility of linking the tensor approach
to the Hilbert-space approach in the treatment of adjustment theory.

An extensive discussion and evaluation of the present geometrical
approach in its relation to the L.S. adjustment has already been included
in Chapter 1. In this final chapter we concentrate on two main items:
1) summarizing, in an explicit form, the most important L.S. results, and
2) inspecting the analogy between the present approach and the Hilbert-
space approach to the L.S. adjustment theory. The first item will be
further subdivided into the parts pertaining to the parametric method,
to the condition method, and to verifications between the two methods.
All the formulas containing either of the weight matrices $P_{\hat{\ell}}$ and $P_{\hat{\varepsilon}}$
(or both) are new and do not appear to have an equivalent in the standard
adjustment literature.

Parametric method. Presented below are several formulas which can be transcribed from the latter part of Section 5.1 or, for the most part, deduced directly from the right-hand side of the diagram in Fig. 3:

$$d\hat{x} = (A^T P A)^{-1} A^T P d\ell \quad , \qquad C_{\hat{x}} = (A^T P A)^{-1} \; ;$$

$$d\hat{\ell} = A(A^T P A)^{-1} A^T P d\ell \; , \qquad C_{\hat{\ell}} = A(A^T P A)^{-1} A^T \; ;$$

$$P_{\hat{\ell}} = P A (A^T P A)^{-1} A^T P \; ;$$

$$\hat{\varepsilon} = d\ell - d\hat{\ell} \; , \qquad C_{\hat{\varepsilon}} = C - C_{\hat{\ell}} \; ;$$

$$P_{\hat{\varepsilon}} = P - P_{\hat{\ell}} \; ;$$

$$d\hat{f} = \tilde{A} d\hat{x} = \tilde{H} d\hat{\ell} \; , \qquad C_{\hat{f}} = \tilde{A} C_{\hat{x}} \tilde{A}^T = \tilde{H} C_{\hat{\ell}} \tilde{H}^T \; ;$$

$$\hat{\varepsilon}^T P \hat{\varepsilon} = d\ell^T P d\ell + d\hat{x}^T u \; , \qquad u = - A^T P d\ell \; .$$

Condition method. The formulas below can be transcribed from the latter part of Section 5.2 or, except for the L.S. quadratic form, deduced directly from the left-hand side of the diagram in Fig. 3:

$$\hat{\varepsilon} = C B^T (B C B^T)^{-1} w \; , \qquad C_{\hat{\varepsilon}} = C B^T (B C B^T)^{-1} B C \; ;$$

$$P_{\hat{\varepsilon}} = B^T (B C B^T)^{-1} B \; ;$$

$$\hat{\varepsilon}^T P \hat{\varepsilon} = w^{*T} w \; , \qquad w^* = (B C B^T)^{-1} w \; .$$

If the initial values ℓ° are not used (the usual procedure in practice), one can express directly the adjusted observations $\hat{\ell}$ from the observed

values $\tilde{\ell}$ and the error estimates $\hat{\varepsilon}$ as

$$\hat{\ell} = \tilde{\ell} - \hat{\varepsilon} \ .$$

Verification between the parametric and condition methods. The first
formula below serves most often as a verification for the parametric
method alone; all of the verification formulas can be transcribed from the
next to the last paragraph in Section 5.2 or, with the exception of ranks,
from the diagram in Fig. 3:

$$A^T P \hat{\varepsilon} = 0 \ ;$$

$$BA = 0 \ ;$$

$$A(A^T PA)^{-1} A^T P + CB^T (BCB^T)^{-1} B = I \ ;$$

$$C_{\hat{\ell}} P_{\hat{\ell}} + C_{\hat{\varepsilon}} P_{\hat{\varepsilon}} = I \ .$$

The rank relationships behave as their matrix counterparts (the latter
have been included in the parametric method above). The rank values
corresponding to the first two lines below appear in the third line:

$$\text{rank } C_{\hat{\ell}} + \text{rank } C_{\hat{\varepsilon}} = \text{rank } C \ ,$$

$$\text{rank } P_{\hat{\ell}} + \text{rank } P_{\hat{\varepsilon}} = \text{rank } P \ ,$$

$$u \quad + \quad r \quad = \quad n \ .$$

Hilbert-space analogy. This part will be related to the Seminar pre-
sented by Vaníček and discussed to some extent in Chapter 1. Since this
Seminar is not yet available in open literature, the references to it will

be stated explicitly. Clearly, one of the most important statements in the Seminar from the point of view of the approach presented herein is, "whether $P_{\mathcal{L}}$ (the weight matrix P) is considered to correspond to covariant or contravariant metric tensor is a more or less arbitrary decision". This has been confirmed via the traditional versus the new identification. The identical outcome demonstrates the perfect duality between the covariant and contravariant "spaces", thereby suggesting that the tensor approach to the L.S. adjustment process can be considered as a classical case of the Hilbert-space approach to the same problem.

We shall complete this statement by identifying the contravariant and covariant tensor spaces with dual Hilbert spaces and describing selected tensors and tensor operations in this context. In so doing, we shall initially proceed along the lines of the Seminar (henceforth abbreviated as [S]) and point out the correspondences with the tensor approach. For example, in [S] the positive-definite matrix P, called weight on Hilbert space $\mathcal{L}$, induces the scalar product

$$<t,t^\circ> = t^T P t^\circ .$$

In the tensor approach (abbreviated as [T]), the Hilbert space $\mathcal{L}$ can be represented by the contravariant space (containing t^r and $t^{\circ r}$) with weight g_{sr}, and the scalar product can be written as

$$<t,t^\circ> = t^s g_{sr} t^{\circ r} .$$

The norm $||.||$ and the distance $d(.,.)$ have the same definition in both [S] and [T], based on the scalar product. As an example in [T], we have

$$d(t,t^\circ) = ||t - t^\circ|| = \sqrt{<t - t^\circ, \ t - t^\circ>} = \sqrt{(t^s - t^{\circ s})g_{sr}(t^r - t^{\circ r})} \ .$$

If $\mathcal{L}^*$ with weight C is the dual Hilbert space to $\mathcal{L}$, we can represent $\mathcal{L}^*$ by the covariant space with weight g^{rs}. In [S], P is the duality operator from $\mathcal{L}$ to $\mathcal{L}^*$, as is g_{sr} in [T]; we have, for example, $g_{sr}t^r = t_s$ (which is indeed in $\mathcal{L}^*$). Furthermore, since $P^{-1} = C$ exists by definition, C in [S] is the inverse duality operator from $\mathcal{L}^*$ to $\mathcal{L}$. The same holds, of course, in [T] with regard to g^{rs}. As in [S] the scalar product, norm, and distance are now defined on the (dual) covariant space $\mathcal{L}^*$ as well; for example, we write

$$<t,t^\circ> = t_r g^{rs} t^\circ_s \ ,$$

which has the same value as the scalar product defined on $\mathcal{L}$. Similar to [S], we can state that $\mathcal{L}^*$ is isometric to $\mathcal{L}$, as well as reflective (the dual of $\mathcal{L}^*$ is $\mathcal{L}$). We also note that either of the contravariant or covariant spaces in [T] could have been associated with $\mathcal{L}$, etc.

In the sequel, we shall no longer present the definitions or derivations in the terminology of [S] with a subsequent interpretation in [T]. The close analogy between these two approaches makes it expedient to proceed directly according to [T] and provide a brief Hilbert-space interpretation as needed. Another reason behind this decision stems from the fact that in certain areas the derivations in [T] have gone beyond the topics treated in [S]. In pursuing [T] Fig. 2 proves helpful. As has been stated earlier, the vector components (boxes in the figure) also represent the pertinent spaces. This is indicated in the following correspondences:

$$dx^r \ \ldots\ldots \ \mathcal{L}, \qquad dx_s \ \ldots\ldots \ \mathcal{L}^* \ ;$$

$$du^\alpha \ \ldots\ldots \ \mathcal{X}, \qquad du_\beta \ \ldots\ldots \ \mathcal{X}^* \ .$$

Here $\mathcal{X}$ is the parameter space of dimensions u, and $\mathcal{X}$, $\mathcal{X}^*$ are defined to have parallel structure to $\mathcal{L}$, $\mathcal{L}^*$. Further we have

$$dx'^r \ \ldots\ldots \ \mathcal{L}', \qquad dx'_s \ \ldots\ldots \ \mathcal{L}^{*'} \ ,$$

where $\mathcal{L}'$ is a subspace of $\mathcal{L}$; $\mathcal{L}^{*'}$ is then the corresponding subspace of $\mathcal{L}^*$.

The tensor formulation in the Hilbert-space context can be adequately illustrated by retracing a few steps in the parametric method. In representing the tensor A_α^r by the matrix A, we confirm the rule mentioned in [S] that while A is the linear operator from $\mathcal{X}$ into $\mathcal{L}'$, A^T is the (inverse) linear operator from $\mathcal{L}^{*'}$ to $\mathcal{X}^*$. In particular, Fig. 2 yields

$$dx'^r = A_\alpha^r du^\alpha \ , \qquad du_\beta = A_\beta^s dx'_s \ .$$

Similarly, we deduce

$$du^\alpha = Q_r^\alpha dx'^r \ , \qquad dx'_s = Q_s^\beta du_\beta \ ,$$

which means that if Q is a linear operator from $\mathcal{L}'$ to $\mathcal{X}$, Q^T is a linear operator from $\mathcal{X}^*$ to $\mathcal{L}^{*'}$. We also confirm that the Hilbertian orthogonal complement $\mathcal{L}'^+$ of $\mathcal{L}'$ in $\mathcal{L}$ is a space of dimensions n-u populated by vectors (in contravariant components) orthogonal to dx'.

Upon introducing the L.S. principle, dx''^r is restricted to $\mathcal{L}'^+$ and dx''_s is similarly restricted to $\mathcal{L}^{*'^+}$, the Hilbertian orthogonal complement

-98-

of $\mathcal{L}^{*'}$ in $\mathcal{L}^{*}$. Accordingly,

$$A_\beta^S dx''_S = 0 \ , \qquad du_\beta = A_\beta^S dx'_S = A_\beta^S dx_S \ ;$$

$$Q_r^\alpha dx''^r = 0 \ , \qquad du^\alpha = Q_r^\alpha dx'^r = Q_r^\alpha dx^r \ .$$

Upon identifying the pertinent operators with matrices, the above shows that the L.S. principle leads to the statement that A^T is the operator not only from $\mathcal{L}^{*'}$ to X^*, but also from $\mathcal{L}^*$ to X^*, and that Q is the operator not only from $\mathcal{L}'$ to X, but also from $\mathcal{L}$ to X.

In continuing in this manner, most of our derivations could have been readily transcribed in the Hilbert-space terminology. Since the present tensor approach has all the ingredients of a classical Hilbert-space approach, the latter can be carried out -- or at least understood -- in terms of the "pictorial" differential geometry with few or no abstract principles. This could help to elucidate the Hilbert-space theory and applications for those who are familiar with a few basic principles of tensor analysis.

ACKNOWLEDGEMENTS

REFERENCES

1. A. MARUSSI: Fondamenti di Geodesia Intrinseca. Publication by the
    Italian Geodetic Commission, No. III 7. Milano, 1951.

2. M. HOTINE: Mathematical Geodesy. Monogr. Ser. Vol. 2, Envir. Sci.
    Serv. Admin., Washington, D.C., 1969.

3. P. VANÍČEK: "Tensor Structure and the Least Squares". Bulletin
    Géodésique, Vol. 53, No. 3, pp. 221-225, 1979.

4. M. MOHAMMAD-KARIM: Diagrammatic Approach to Solve Least-Squares
    Adjustment and Collocation Problems. Technical Report No. 83,
    Department of Surveying Engineering, University of New Brunswick,
    Fredericton, N.B., Canada, September 1981.

5. J. M. TIENSTRA: "An Extension of the Technique of the Methods of
    Least Squares to Correlated Observations". Bulletin Géodés e,
    No. 6, 1947.

6. J. M. TIENSTRA: "The Foundation of the Calculus of Observatic
    and the Method of Least Squares". Bulletin Géodésique, No. ,
    1948.

7. J. M. TIENSTRA: Theory of the Adjustment of Normally Distributed
    Observations. Edited by his friends. N. V. Uitgeverij,
    Argus, Amsterdam, 1956.

8. W. BAARDA: Statistical Concepts in Geodesy. Publications on Geodesy
    by the Netherlands Geodetic Commission, New Series, Vol. 2,
    No. 4, Delft, 1967.

9. W. BAARDA: Testing Procedure for Use in Geodetic Networks. Publica-
    tions on Geodesy by the Netherlands Geodetic Commission,
    New Series, Vol. 2, No. 5, Delft, 1968.

10. W. BAARDA: Vereffeningstheorie, Vol. 1 & 2. Computing Centre of
    the Delft Geodetic Institute, Technische Hogeschool, Delft,
    1967 & 1970.

11. A. H. KOOIMANS: "Principles of the Calculus of Observations".
    Rapport Spécial, Neuvième Congrès International des Géomètres,
    Pays-Bas, pp. 301-310, 1958.

12. W. M. KAULA: Theory of Satellite Geodesy. Blaisdell Publishing Co.,
    Waltham, Mass., 1966.

13. H. SCHEFFÉ: The Analysis of Variance. John Wiley & Sons, Inc.,
    New York, N.Y., 1967.

ATE
LME